

Oligomer search for  
length of 6

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2006, 09:10:59 ; Search time 27446 Seconds  
(without alignments)  
17665.574 Million cell updates/sec

Title: US-10-717-580-11

Perfect score: 7582

Sequence: 1 caacaatcggtataaacc.....tattaaattctgcactgtg 7582

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 6366136 seqs, 31973710525 residues

Word size : 6

Total number of hits satisfying chosen parameters: 12690550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 90 summaries

Database :

GenEmbl.\*

1: gb\_env.\*

2: gb\_pat.\*

3: gb\_ph.\*

4: gb\_pl.\*

5: gb\_pr.\*

6: gb\_ro.\*

7: gb\_sts.\*

8: gb\_sy.\*

9: gb\_un.\*

10: gb\_vl.\*

11: gb\_ov.\*

12: gb\_htg.\*

13: gb\_in.\*

14: gb\_om.\*

15: gb\_ba.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7582	100.0	7582	2	BD196245 Endogenet
2	7582	100.0	7582	2	BD267487 Endogenet
3	7582	100.0	7582	2	AX000966 Sequence
4	7582	100.0	7582	2	AX027480 Sequence
5	888	11.7	2372	2	BD196244 Endogenet
6	888	11.7	2372	2	BD267486 Endogenet
7	888	11.7	2372	2	AX000965 Sequence
8	888	11.7	2372	2	AX027479 Sequence
9	888	11.7	2372	5	AF072504 Homo sapi
10	779	10.3	2599	2	BD231806 Nucleic s
11	779	10.3	2599	2	AR699755 Sequence
12	779	10.3	2599	2	AR007978 Sequence
13	779	10.3	8523	2	AR612522 Sequence
14	779	10.3	10222	5	AY101582 Homo sapi
15	779	10.3	10222	5	AY101585 Homo sapi
16	779	10.3	10499	2	BD221808 Nucleic s
17	779	10.3	10499	2	AR699757 Sequence
18	779	10.3	10499	2	AX007980 Sequence

19	779	10.3	5693	2	AX329572	AX329572 Sequence
20	779	10.3	5693	5	HSAC000064	AC000064 Human BAC
c 21	779	10.3	149194	5	AC007566	AC007566 Homo sapi
22	731	9.6	1617	5	AF156963	AF156963 Homo sapi
23	731	9.6	1948	2	BD196239	BD196239 Endogenet
24	731	9.6	1948	2	BD267481	BD267481 Endogenet
25	731	9.6	1948	2	AX000950	AX000950 Sequence
26	731	9.6	1948	2	AX027474	AX027474 Sequence
27	731	9.6	1948	5	AF072505	AF072505 Homo sapi
28	731	9.6	2055	2	BD221827	BD221827 Nucleic s
29	731	9.6	2055	2	AR699776	AR699776 Sequence
30	731	9.6	2055	2	AX007999	AX007999 Sequence
31	731	9.6	2694	5	AF520483S2	AF520484 Homo sapi
32	731	9.6	2694	5	AF520485S2	AF520486 Homo sapi
33	731	9.6	2694	5	AF520493S2	AF520494 Homo sapi
34	731	9.6	2694	5	AF520495S2	AF520496 Homo sapi
35	731	9.6	2694	5	AF520497S2	AF520498 Homo sapi
36	731	9.6	2694	5	AF520499S2	AF520500 Homo sapi
37	731	9.6	2694	5	AF520501S2	AF520502 Homo sapi
38	731	9.6	2694	5	AF520503S2	AF520504 Homo sapi
39	731	9.6	2694	5	AF520507S2	AF520508 Homo sapi
40	731	9.6	2694	5	AF520509S2	AF520510 Homo sapi
41	731	9.6	2694	5	AF520511S2	AF520512 Homo sapi
42	731	9.6	2694	5	AF520513S2	AF520514 Homo sapi
43	731	9.6	2694	5	AF520515S2	AF520516 Homo sapi
44	731	9.6	2694	5	AF520523S2	AF520524 Homo sapi
45	731	9.6	2694	5	AF520525S2	AF520526 Homo sapi
46	731	9.6	2694	5	AF520527S2	AF520528 Homo sapi
47	731	9.6	2694	5	AF520531S2	AF520532 Homo sapi
48	731	9.6	2694	5	AF520533S2	AF520534 Homo sapi
49	731	9.6	2694	5	AF520535S2	AF520536 Homo sapi
50	731	9.6	2694	5	AF520537S2	AF520538 Homo sapi
51	731	9.6	2694	5	AF520539S2	AF520540 Homo sapi
52	731	9.6	2694	5	AF520541S2	AF520542 Homo sapi
53	731	9.6	2694	5	AF520543S2	AF520544 Homo sapi
54	731	9.6	2694	5	AF520545S2	AF520546 Homo sapi
55	731	9.6	2694	5	AF520547S2	AF520548 Homo sapi
56	731	9.6	2694	5	AF520549S2	AF520550 Homo sapi
57	731	9.6	2694	5	AF520551S2	AF520552 Homo sapi
58	731	9.6	2694	5	AF520553S2	AF520554 Homo sapi
59	731	9.6	2694	5	AF520555S2	AF520556 Homo sapi
60	731	9.6	2694	5	AF520557S2	AF520558 Homo sapi
61	731	9.6	2782	2	BD196241	BD196241 Endogenet
62	731	9.6	2782	2	BD267483	BD267483 Endogenet
63	731	9.6	2782	2	AX000962	AX000962 Sequence
64	731	9.6	2782	2	AX027476	AX027476 Sequence
65	731	9.6	2782	2	AX133396	AX133396 Sequence
66	728	9.6	2930	2	AX355872	AX355872 Sequence
67	728	9.6	2930	5	AF208161	AF208161 Homo sapi
68	728	9.6	10222	5	AY101583	AY101583 Homo sapi
69	724	9.5	2938	2	BD196236	BD196236 Endogenet
70	724	9.5	2938	2	BD267478	BD267478 Endogenet
71	724	9.5	2938	2	AX000957	AX000957 Sequence
72	724	9.5	2938	2	AX027471	AX027471 Sequence
73	724	9.5	2938	5	AF072499	AF072499 Homo sapi
74	680	9.0	1624	5	AF506835	AF506835 Homo sapi
75	680	9.0	1860	5	AF513360	AF513360 Homo sapi
76	680	9.0	2694	5	AF520477S2	AF520478 Homo sapi
77	680	9.0	2694	5	AF520479S2	AF520480 Homo sapi
78	680	9.0	2694	5	AF520481S2	AF520482 Homo sapi
79	680	9.0	2694	5	AF520487S2	AF520488 Homo sapi
80	680	9.0	2694	5	AF520489S2	AF520490 Homo sapi
81	680	9.0	2694	5	AF520505S2	AF520506 Homo sapi
82	680	9.0	2694	5	AF520517S2	AF520518 Homo sapi
83	680	9.0	2694	5	AF520519S2	AF520520 Homo sapi
84	680	9.0	2694	5	AF520521S2	AF520522 Homo sapi
85	680	9.0	2694	5	AF520529S2	AF520530 Homo sapi
86	680	9.0	2694	5	AF520532S2	AF520533 Homo sapi
87	680	9.0	2694	5	AF520545S2	AF520546 Homo sapi
88	680	9.0	2694	5	AF520559S2	AF520560 Homo sapi
89	680	9.0	2946	2	AK177269	AK177269 Sequence
90	680	9.0	2946	2	BD086133	BD086133 Secreted

## ALIGNMENTS

```

RESULT 1
BD196245
LOCUS      BD196245
DEFINITION Endogenous retroviral sequences, associated with autoimmune diseases and/or with pregnancy disorders.
ACCESSION  BD196245
VERSION    BD196245.1
KEYWORDS   JP 2002512530-A/11.
SOURCE     unidentified
ORGANISM   unidentified
            unclassified sequences.
REFERENCE  1 (bases 1 to 7582)
AUTHORS   Beseme, F., Blond, J.L., Bouton, O., Mandrand, B. and Mallet, F.
TITLE     Endogenous retroviral sequences, associated with autoimmune diseases and/or with pregnancy disorders
JOURNAL   Patent: JP 2002512530-A 11 23-APR-2002;
          BIO MERIEUX
COMMENT   OS Unidentified
          PN JP 2002512530-A/11
          PD 23-APR-2002
          PF 06-JUL-1998 JP 1999508244
          PR 07-JUL-1997 FR 97/08815
          PI FREDERIC BESEME, JEAN LUC BLOND, OLIVIER BOUTON, BERNARD MANDRAND,
          PI FRANCOIS MALLET
          PC C12N15/48, C07K14/15, C12Q1/68, C07K16/10, G01N33/569 CC
          Strandedness: Single;
          CC Topology: Linear;
          CC Endogenous retroviral sequences, associated with autoimmune diseases
          CC and/or with pregnancy disorders
          FH Key Location/Qualifiers
          FT source 1..7582
          /organism="Unidentified".
          /Location/Qualifiers
          1..7582
          /organism="unidentified"
          /mol_type="genomic DNA"
          /db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 7582; DB 2; Length 7582;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAACAATCGGGATATAAACCAGGCGATTGAGCTGGCAACAGCAGCCCCCTTTGGGTCC 60
DB 1 CAACAATCGGGATATAAACCAGGCGATTGAGCTGGCAACAGCAGCCCCCTTTGGGTCC 60

QY 61 CTTCCCTTTGATGGGAGCTGTTTTCATGCTATTTCCACTCTATTAAATCTGCAACTGCA 120
DB 61 CTTCCCTTTGATGGGAGCTGTTTTCATGCTATTTCCACTCTATTAAATCTGCAACTGCA 120

QY 121 CTCCTTCGTGTCATGTTTCTTAGCGCTCGAGCTGAGCTTTTGTCTCAACCGTCCACCACTGC 180
DB 121 CTCCTTCGTGTCATGTTTCTTAGCGCTCGAGCTGAGCTTTTGTCTCAACCGTCCACCACTGC 180

QY 181 TGTTCGCAACCCGCGCANACTGCGCTGACTCCCATGCCCTCTGGATTCCTGCAAGGTGTC 240
DB 181 TGTTCGCAACCCGCGCANACTGCGCTGACTCCCATGCCCTCTGGATTCCTGCAAGGTGTC 240

QY 241 CGCTGTGCTCCTGATCCAGGARGCGCCATTGCGCTCCCAATTCGGCTTAAAGGCTTGC 300
DB 241 CGCTGTGCTCCTGATCCAGGARGCGCCATTGCGCTCCCAATTCGGCTTAAAGGCTTGC 300

QY 301 CATTTGNCCTGCAACCGCTTAAGTGCCTGGGTTTGTCTTAATTTGAGCTGAACACTTACCT 360
DB 301 CATTTGNCCTGCAACCGCTTAAGTGCCTGGGTTTGTCTTAATTTGAGCTGAACACTTACCT 360

```

```

QY 361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATATAACACTTACCACA 420
DB 361 GGGTTCCATGGTTCTCTTCTGTGACCCACGGCTTCTAATAKAACTATATAACACTTACCACA 420

QY 421 TGGCCCAAGATTCCATTCCTTTGGAATCCGTGAGGSCAAACGAACTCCAGGTCAAGAAATAC 480
DB 421 TGGCCCAAGATTCCATTCCTTTGGAATCCGTGAGGSCAAACGAACTCCAGGTCAAGAAATAC 480

QY 481 GARGCTTCCACCATCTTTGGAAGCGGCTGCTACCTCTTGGAAAGTGGTTTCAACCACTC 540
DB 481 GARGCTTCCACCATCTTTGGAAGCGGCTGCTACCTCTTGGAAAGTGGTTTCAACCACTC 540

QY 541 TTGGGAGCTCTGTGAGCAAGGACCCCGGTACATTTTGGCRAACMSRACGACATCC 600
DB 541 TTGGGAGCTCTGTGAGCAAGGACCCCGGTACATTTTGGCRAACMSRACGACATCC 600

QY 601 MAAGTGATGGGAAACGTTTCCCGCAAGACAAACACGCCCCCTTAAGACGTAATTTCTGGARAAT 660
DB 601 MAAGTGATGGGAAACGTTTCCCGCAAGACAAACACGCCCCCTTAAGACGTAATTTCTGGARAAT 660

QY 661 TGGGAMCAATTTGACCCCTCAGACACTAAGAAAGAAACGACTTATATTTCTTCTGCACTGCC 720
DB 661 TGGGAMCAATTTGACCCCTCAGACACTAAGAAAGAAACGACTTATATTTCTTCTGCACTGCC 720

QY 721 GCCTGGCACTCTGTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACVCTTTTGG 780
DB 721 GCCTGGCACTCTGTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACVCTTTTGG 780

QY 781 TAGAAAAGGCAAAATGGAGTGAAGTGCCATAAGTACAAAATTTCTTTTCAATTAAGAGACAA 840
DB 781 TAGAAAAGGCAAAATGGAGTGAAGTGCCATAAGTACAAAATTTCTTTTCAATTAAGAGACAA 840

QY 841 CTCACAATTAATGTAAGAAAGTGTGATTTATGCCCTCAGAGAACCTTTCAGAGCTACCTCC 900
DB 841 CTCACAATTAATGTAAGAAAGTGTGATTTATGCCCTCAGAGAACCTTTCAGAGCTACCTCC 900

QY 901 CTATCCAGCATCCCGACTCTTCCCGCCTTCCCGCCTTCCCGCCTTCCCGCCTTCCCGCCTTCC 960
DB 901 CTATCCAGCATCCCGACTCTTCCCGCCTTCCCGCCTTCCCGCCTTCCCGCCTTCCCGCCTTCC 960

QY 961 TCCAAAAGGAGATAGACAAAGGTTAAACAGTGAACCAAGAGTGCCTAATTTTCCCAAT 1020
DB 961 TCCAAAAGGAGATAGACAAAGGTTAAACAGTGAACCAAGAGTGCCTAATTTTCCCAAT 1020

QY 1021 TATGACCCCTTCCCAAGAGTGGGAGAGAGAAATTTGGCCCGCAGCAGAGTGCATGCTC 1080
DB 1021 TATGACCCCTTCCCAAGAGTGGGAGAGAGAAATTTGGCCCGCAGCAGAGTGCATGCTC 1080

QY 1081 TTTTCTTCCAGACTTAAAGCAAAATAAAACAGACTTAAAGTAAATTTCTCAGATAATCCT 1140
DB 1081 TTTTCTTCCAGACTTAAAGCAAAATAAAACAGACTTAAAGTAAATTTCTCAGATAATCCT 1140

QY 1141 GATGCTATATTTGRTGTTTCAAGGGTTAGGACAAATTTCTTGTATCTGACATGGAGAGAT 1200
DB 1141 GATGCTATATTTGRTGTTTCAAGGGTTAGGACAAATTTCTTGTATCTGACATGGAGAGAT 1200

QY 1201 ATATATGTCATGCTTAAATCAGACACTAAACCCCAATGAGAGAGTGCACCACTATGTC 1260
DB 1201 ATATATGTCATGCTTAAATCAGACACTAAACCCCAATGAGAGAGTGCACCACTATGTC 1260

QY 1261 AGCCTGAGGTTTGGGATCTCTGGTATCTCAGTCAAGTCAATGGATANGATGACAAACA 1320
DB 1261 AGCCTGAGGTTTGGGATCTCTGGTATCTCAGTCAAGTCAATGGATANGATGACAAACA 1320

QY 1321 GAAGGAAAGANAATGATTTCCCAAGGCGCAGCAGTTCAGCTTASACCTCATTTG 1380
DB 1321 GAAGGAAAGANAATGATTTCCCAAGGCGCAGCAGTTCAGCTTASACCTCATTTG 1380

QY 1381 GGGACACAGAAATCAGTAACATGGGAGATTTGCTGCAAGATTTGCTTAACTTGTGTC 1440
DB 1381 GGGACACAGAAATCAGTAACATGGGAGATTTGCTGCAAGATTTGCTTAACTTGTGTC 1440

QY 1441 TASAAGGACTAAGGAAACTASAGAGAAATCTAAGAAATTAATCTCAATGATGTCCACATA 1500

```





3661 GTGAAYRGATACCTCACTGCTCTAAAGGAGACTTGTGGCTGTGACAGCAACVGTCTTAA 3720  
Db |||||  
3661 GTGAAYRGATACCTCACTGCTCTAAAGGAGACTTGTGGCTGTGACAGCAACVGTCTTAA 3720  
Qy |||||  
3721 TRTCAGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGACCTTGTGCAACTCTTAAC 3780  
Db |||||  
3721 TRTCAGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGACCTTGTGCAACTCTTAAC 3780  
Qy |||||  
3781 CCAGTCNCAATTTCTTCAGACAAATGAAGAAAGATARAAYATACTGTCAACAARTTAAT 3840  
Db |||||  
3781 CCAGTCNCAATTTCTTCAGACAAATGAAGAAAGATARAAYATACTGTCAACAARTTAAT 3840  
Qy |||||  
3841 TCTCAAAACCTAATGCCACTCGAGGACCTTCTTAGAGTTTCCCTTGCACCTGATCCYACCTT 3900  
Db |||||  
3841 TCTCAAAACCTAATGCCACTCGAGGACCTTCTTAGAGTTTCCCTTGCACCTGATCCYACCTT 3900  
Qy |||||  
3901 CAACTTGTATCTAGTGAAGTTCCTTTGTAGAAAAGGACTTCGAAAAGYGGGTATGC 3960  
Db |||||  
3901 CAACTTGTATCTAGTGAAGTTCCTTTGTAGAAAAGGACTTCGAAAAGYGGGTATGC 3960  
Qy |||||  
3961 AGTGTCTAGTGAATATGGAATATTTGAAGTAAATCCCCTCACTCCAGGAACTAGTGTCTYA 4020  
Db |||||  
3961 AGTGTCTAGTGAATATGGAATATTTGAAGTAAATCCCCTCACTCCAGGAACTAGTGTCTYA 4020  
Qy |||||  
4021 GCTRCAGAACTAATAGCCCTCACTKGGGCACTAGAAATTTAGGAGAAAGAAAGGAYAA 4080  
Db |||||  
4021 GCTRCAGAACTAATAGCCCTCACTKGGGCACTAGAAATTTAGGAGAAAGAAAGGAYAA 4080  
Qy |||||  
4081 TATATATACAGACTCTRTATATGCTTACCTAGTCTNTCCATGCCATGMRGCAATATGSAR 4140  
Db |||||  
4081 TATATATACAGACTCTRTATATGCTTACCTAGTCTNTCCATGCCATGMRGCAATATGSAR 4140  
Qy |||||  
4141 AGAAAGGGAATTCCTAACTTCYAGRGAAACACTATCAWACATCAGGAAGCCATTAGGAR 4200  
Db |||||  
4141 AGAAAGGGAATTCCTAACTTCYAGRGAAACACTATCAWACATCAGGAAGCCATTAGGAR 4200  
Qy |||||  
4201 ATTATATYGGGWTACAGAAACCTARAGAGTGGAGTCTTACACTGCGGGTCACTCA 4260  
Db |||||  
4201 ATTATATYGGGWTACAGAAACCTARAGAGTGGAGTCTTACACTGCGGGTCACTCA 4260  
Qy |||||  
4261 NAAAGGAAGRAAGGGAATASAGRGAAATGCAAGCAKATATGAAGCMAAAGAGC 4320  
Db |||||  
4261 NAAAGGAAGRAAGGGAATASAGRGAAATGCAAGCAKATATGAAGCMAAAGAGC 4320  
Qy |||||  
4321 TGCAAGGAGGACCCCTCCATTAGAAATGCTTATTAACCTTCCCTTAGTATAGGGTAATCC 4380  
Db |||||  
4321 TGCAAGGAGGACCCCTCCATTAGAAATGCTTATTAACCTTCCCTTAGTATAGGGTAATCC 4380  
Qy |||||  
4381 CTTCCGGGAACCAAGCCAGTACTCAGCAGGAGAAACAGAAATGGGNAACCTCAGGAG 4440  
Db |||||  
4381 CTTCCGGGAACCAAGCCAGTACTCAGCAGGAGAAACAGAAATGGGNAACCTCAGGAG 4440  
Qy |||||  
4441 CAGTTTTCTCCCTCGGACGGTTAGCCACTGAAGAGGGAATACTTTTGCCTGCAAC 4500  
Db |||||  
4441 CAGTTTTCTCCCTCGGACGGTTAGCCACTGAAGAGGGAATACTTTTGCCTGCAAC 4500  
Qy |||||  
4501 TATCCAAATGGAATTAATTAACCCCTTCAATCAAACTTTCACTTAGGCATCGATAGCAC 4560  
Db |||||  
4501 TATCCAAATGGAATTAATTAACCCCTTCAATCAAACTTTCACTTAGGCATCGATAGCAC 4560  
Qy |||||  
4561 CCATCATATGGCCAAATGATTTTACTGACAGGCTTTTCAAACTATCAAGCARAT 4620  
Db |||||  
4561 CCATCATATGGCCAAATGATTTTACTGACAGGCTTTTCAAACTATCAAGCARAT 4620  
Qy |||||  
4621 AKTCAGGGCTGTGAATGTGCCARAAATATCCCTGCTCTATCGCAAGCTCCTTC 4680  
Db |||||  
4621 AKTCAGGGCTGTGAATGTGCCARAAATATCCCTGCTCTATCGCAAGCTCCTTC 4680  
Qy |||||  
4681 AGGAACAAABAACAGGCCATTACCTGGRARAARACTGGCAACTGATTTTACCCACAAG 4740  
Db |||||  
4681 AGGAACAAABAACAGGCCATTACCTGGRARAARACTGGCAACTGATTTTACCCACAAG 4740

4741 CCCAAACCTCAGGATTTTCAGTATCTACTAGTCTGGGTARATACCTTTTACCGGTTGGCA 4800  
Db |||||  
4741 CCCAAACCTCAGGATTTTCAGTATCTACTAGTCTGGGTARATACCTTTTACCGGTTGGCA 4800  
Qy |||||  
4801 RAGGCTTCCCTCTGAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCAATGAA 4860  
Db |||||  
4801 RAGGCTTCCCTCTGAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCAATGAA 4860  
Qy |||||  
4861 ATAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTTTCAG 4920  
Db |||||  
4861 ATAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTTTCAG 4920  
Qy |||||  
4921 GCCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACCTCGGCC 4980  
Db |||||  
4921 GCCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACCTCGGCC 4980  
Qy |||||  
4981 TGAAGGCCACAGTCTCTCAGGGAAGGTGAGAAAATGAATGAATAAAYACTCAAGGAATCTA 5040  
Db |||||  
4981 TGAAGGCCACAGTCTCTCAGGGAAGGTGAGAAAATGAATGAATAAAYACTCAAGGAATCTA 5040  
Qy |||||  
5041 AAAAAGCAAAACCCAGGAAACCCACCTCACATGGCTGTCTGTGCTTATAGCCTTAAA 5100  
Db |||||  
5041 AAAAAGCAAAACCCAGGAAACCCACCTCACATGGCTGTCTGTGCTTATAGCCTTAAA 5100  
Qy |||||  
5101 AGAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAG 5160  
Db |||||  
5101 AGAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAG 5160  
Qy |||||  
5161 CCCTTCAATAACCAATGACTTGTGACCCCAAGACAGCCAACTTAGTTGACAGATCA 5220  
Db |||||  
5161 CCCTTCAATAACCAATGACTTGTGACCCCAAGACAGCCAACTTAGTTGACAGATCA 5220  
Qy |||||  
5221 CCTCTTAGCCAAATATCAACAGTTCTTAAACATTTAAGGAACCTTATCCCTGAGAG 5280  
Db |||||  
5221 CCTCTTAGCCAAATATCAACAGTTCTTAAACATTTAAGGAACCTTATCCCTGAGAG 5280  
Qy |||||  
5281 AGGAAAAGAACTATTCCACCCWGTGACATGTTAGTCAAGTCCCTTCTCTTAAT 5340  
Db |||||  
5281 AGGAAAAGAACTATTCCACCCWGTGACATGTTAGTCAAGTCCCTTCTCTTAAT 5340  
Qy |||||  
5341 CCCATCCTTAGATACATCTCTGGGAAGGACCTTACCCAGTCACTTTATACCCCACTG 5400  
Db |||||  
5341 CCCATCCTTAGATACATCTCTGGGAAGGACCTTACCCAGTCACTTTATACCCCACTG 5400  
Qy |||||  
5401 CGGTTAAAGTGGCTGGAGTCTTGATACATCACTTTGAGTCAAACTCTGGATAC 5460  
Db |||||  
5401 CGGTTAAAGTGGCTGGAGTCTTGATACATCACTTTGAGTCAAACTCTGGATAC 5460  
Qy |||||  
5461 TGCCAAAGGAACCTGAAATCCAGGAGCAACGCTAGTCTTCTCTGTGAACCTCTAGAGG 5520  
Db |||||  
5461 TGCCAAAGGAACCTGAAATCCAGGAGCAACGCTAGTCTTCTCTGTGAACCTCTAGAGG 5520  
Qy |||||  
5521 ATTTGCGCTCTCTTCAAAACCAACAGGAGGAAGTAATAAATCATAAATCCCTCC 5580  
Db |||||  
5521 ATTTGCGCTCTCTTCAAAACCAACAGGAGGAAGTAATAAATCATAAATCCCTCC 5580  
Qy |||||  
5581 ATGSCCTCTCTTATCATATTTTCTCTKTASTGTSTTTTACCTCTTCTCACTCTCACT 5640  
Db |||||  
5581 ATGSCCTCTCTTATCATATTTTCTCTKTASTGTSTTTTACCTCTTCTCACTCTCACT 5640  
Qy |||||  
5641 GCACCCCTCATGCGCTGTATGACCACTAGTCTCCCTYACCMAGAGTTTCTATGGAGA 5700  
Db |||||  
5641 GCACCCCTCATGCGCTGTATGACCACTAGTCTCCCTYACCMAGAGTTTCTATGGAGA 5700  
Qy |||||  
5701 ATGACGCTCCCGGAATATTTGATGCCCATCTGATAGGAGTCTTTTAAAGGAACCCCT 5760  
Db |||||  
5701 ATGACGCTCCCGGAATATTTGATGCCCATCTGATAGGAGTCTTTTAAAGGAACCCCT 5760  
Qy |||||  
5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCGCACTCTTTGCA 5820  
Db |||||  
5761 ACCTTCACTGCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCGCACTCTTTGCA 5820  
Qy |||||  
5821 CATGCAAAATCTCATTATTGGACAGGAATAATGATTAATCTCTAGTTGTCTCTGAGGACTT 5880

Db 5821 CATGCAAACTACTATTAATGGACAGGAAATGATTAATCCTAGTTGTCCTGGAGACTT 5880.  
 QY 5881 GAGTCACTGCTGTTGGACTTACTTCAACCAAACTGGTATGTCGATGGGGTGGAGTT 5940  
 Db 5881 GAGTCACTGCTGTTGGACTTACTTCAACCAAACTGGTATGTCGATGGGGTGGAGTT 5940  
 QY 5941 CAAGATCAGGCAAGAGAAAACATGTAAGAAAGTAACTCTCCCAACTCAACCSGGGTACAT 6000  
 Db 5941 CAAGATCAGGCAAGAGAAAACATGTAAGAAAGTAACTCTCCCAACTCAACCSGGGTACAT 6000  
 QY 6001 GGCACTCTAGCCCTCAAGAGGACTAGATCTCTCAAAAATACATGAAGAAACCTTCCGTACC 6060  
 Db 6001 GGCACTCTAGCCCTCAAGAGGACTAGATCTCTCAAAAATACATGAAGAAACCTTCCGTACC 6060  
 QY 6061 CATACTCGCTGTAAGCCTATTATTAATACACCCCTCACTGGGCTCCATGAGGTCTCGGCC 6120  
 Db 6061 CATACTCGCTGTAAGCCTATTATTAATACACCCCTCACTGGGCTCCATGAGGTCTCGGCC 6120  
 QY 6121 CAAACCCCTACTAACTGTTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTCAATC 6180  
 Db 6121 CAAACCCCTACTAACTGTTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTCAATC 6180  
 QY 6181 CTTGTACCTGAACAATGGAACTTCAAGACAGAAATAAACAACACTTCCGTTTGTAGTA 6240  
 Db 6181 CTTGTACCTGAACAATGGAACTTCAAGACAGAAATAAACAACACTTCCGTTTGTAGTA 6240  
 QY 6241 GGNACCTCTGTTTCCAAATSTGAAATTAACCCATACCTCAAACTCACTGTGTAAATTT 6300  
 Db 6241 GGNACCTCTGTTTCCAAATSTGAAATTAACCCATACCTCAAACTCACTGTGTAAATTT 6300  
 QY 6301 AGCAATACTACATACACCAACCTCCCAATGATCAGTGGTGGTAACTCTCCACACAA 6360  
 Db 6301 AGCAATACTACATACACCAACCTCCCAATGATCAGTGGTGGTAACTCTCCACACAA 6360  
 QY 6361 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTAGCTCAGCCCTATGTTGTTTG 6420  
 Db 6361 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTAGCTCAGCCCTATGTTGTTTG 6420  
 QY 6421 AATGGCTCTCAGAACTATGCTTCTCTCAATCTTAGTGCCCCCYATGCCACTCTAC 6480  
 Db 6421 AATGGCTCTCAGAACTATGCTTCTCTCAATCTTAGTGCCCCCYATGCCACTCTAC 6480  
 QY 6481 ACTGAACAGATTTATACAGTTATGTCATATCTAGCCCGGCAACAAAGAGTACCCATT 6540  
 Db 6481 ACTGAACAGATTTATACAGTTATGTCATATCTAGCCCGGCAACAAAGAGTACCCATT 6540  
 QY 6541 CTTCTCTTTGTTATAGGACGAGGTGCTAGGTGCACTAGGTACTGGCAATTCGGCGTATC 6600  
 Db 6541 CTTCTCTTTGTTATAGGACGAGGTGCTAGGTGCACTAGGTACTGGCAATTCGGCGTATC 6600  
 QY 6601 ACAACCTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG 6660  
 Db 6601 ACAACCTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG 6660  
 QY 6661 GTCCGCGACTCCTGCTGCTACCTTGAAGATCACTTAACTCCCTAGCAGCAGTACCTTT 6720  
 Db 6661 GTCCGCGACTCCTGCTGCTACCTTGAAGATCACTTAACTCCCTAGCAGCAGTACCTTT 6720  
 QY 6721 CCAAAATCGAAGAGCTTACACTGCTTAACTCAATCGGAATCGTCACTGAGAAAGTTAAAGAA 6780  
 Db 6721 CCAAAATCGAAGAGCTTACACTGCTTAACTCAATCGGAATCGTCACTGAGAAAGTTAAAGAA 6780  
 QY 6781 GGGGAGAAATGCTGTTATTATGTTAACTCAATCGGAATCGTCACTGAGAAAGTTAAAGAA 6840  
 Db 6781 GGGGAGAAATGCTGTTATTATGTTAACTCAATCGGAATCGTCACTGAGAAAGTTAAAGAA 6840  
 QY 6841 ATTCSAGATCGAATACAACTAGCAGAGAGCTTCGAAACACTGGACCTCGGGGCTC 6900  
 Db 6841 ATTCSAGATCGAATACAACTAGCAGAGAGCTTCGAAACACTGGACCTCGGGGCTC 6900  
 QY 6901 CTCAGGCRATGATCCCTGGATTCTCCCTCTTAGGACCTCTAGCAGCTATATATTG 6960

Db 6901 CTCAGCRRATGGATGCCCTGGATTTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATTG 6960  
 QY 6961 CTACTCTCTTTTGGACCTCTGTATCTTTTACCTCTCTTTTAACTTTCTCTCTCCAGATC 7020  
 Db 6961 CTACTCTCTTTTGGACCTCTGTATCTTTTACCTCTCTTTTAACTTTCTCTCTCCAGATC 7020  
 QY 7021 GAAGCTGTAAACTCAAAATGGAGCCCAAGATGCACTCAAGACTAAGATCTTACCCGAGA 7080  
 Db 7021 GAAGCTGTAAACTCAAAATGGAGCCCAAGATGCACTCAAGACTAAGATCTTACCCGAGA 7080  
 QY 7081 CCCCTGGACCGGCTGTAGCCCAAGATCTGTATGTTTAACTCAAGACTAAGATCTTACCC 7140  
 Db 7081 CCCCTGGACCGGCTGTAGCCCAAGATCTGTATGTTTAACTCAAGACTAAGATCTTACCC 7140  
 QY 7141 GAGGAAATCTAGCTGCAACAACTCTATACCCCAATTCAGCAGGAGCAGTGTAGAGC 7200  
 Db 7141 GAGGAAATCTAGCTGCAACAACTCTATACCCCAATTCAGCAGGAGCAGTGTAGAGC 7200  
 QY 7201 GGTSGTGGCCCAACCTCCCAACAGCAGTGTAGGTTTCTGTTGAGATGGGAGCTGAGA 7260  
 Db 7201 GGTSGTGGCCCAACCTCCCAACAGCAGTGTAGGTTTCTGTTGAGATGGGAGCTGAGA 7260  
 QY 7261 GACAGACTAGCTGCAATTTCTAGCTGATTAAGAAATCCYTAAGCCTAGTGTGGAGGTTG 7320  
 Db 7261 GACAGACTAGCTGCAATTTCTAGCTGATTAAGAAATCCYTAAGCCTAGTGTGGAGGTTG 7320  
 QY 7321 ACCACATCCACTTTTAAACACGGGCTTGCACACTTAGYTACACCTGACCAATTCAGAG 7380  
 Db 7321 ACCACATCCACTTTTAAACACGGGCTTGCACACTTAGYTACACCTGACCAATTCAGAG 7380  
 QY 7381 CTCACATAAATGCTAATTTAGGCAAGACAGGAGGTAAGAAATAGCAATCATYTATTC 7440  
 Db 7381 CTCACATAAATGCTAATTTAGGCAAGACAGGAGGTAAGAAATAGCAATCATYTATTC 7440  
 QY 7441 MTGAGAGCAGCAGCAGGAGGCAATGATCGGATATAACCCCAAGTTCGAGCCGCGAA 7500  
 Db 7441 MTGAGAGCAGCAGGAGGCAATGATCGGATATAACCCCAAGTTCGAGCCGCGAA 7500  
 QY 7501 CGGCAACCCCTTTGGGTCCTCTCTGTTGATGGAGCTCTGTTTCAATGCTATTTCAC 7560  
 Db 7501 CGGCAACCCCTTTGGGTCCTCTCTGTTGATGGAGCTCTGTTTCAATGCTATTTCAC 7560  
 QY 7561 TCTATTAAATCTTGCARCTGCR 7582  
 Db 7561 TCTATTAAATCTTGCARCTGCR 7582

RESULT 2  
 BD267487  
 LOCUS

DEFINITION  
 Endogenous nucleic fragment associated with an autoimmune disease,  
 marking method and reagent.

ACCESSION  
 BD267487

VERSION  
 BD267487.1 GI:33077255

KEYWORDS  
 JP 2002534980-A/28.

SOURCE  
 Homo sapiens (human)

ORGANISM  
 Homo sapiens

REFERENCE  
 1 (bases 1 to 7582)

AUTHORS  
 Baccala,G.P., Mallet,F. and Voisset,C.

TITLE  
 Endogenous nucleic fragment associated with an autoimmune disease,  
 marking method and reagent

JOURNAL  
 Patent: JP 2002534980-A 28 22-OCT-2002;  
 BIO MERIEUX

COMMENT  
 OS Homo sapiens (human)  
 PN JP 2002534980-A/28  
 PD 22-OCT-2002  
 PF 21-JAN-2000 JP 2000594929  
 PR 21-JAN-1999 FR 99/00888  
 PI GLAUCIA PARANHOS BACCALA, FRANCOIS MALLET, CECILE VOISSET PC  
 C12N15/09,C07K14/15,C12Q1/02,C12Q1/68,G01N33/53, G01N33/53, PC

Not proof

7582 bp DNA linear PAT 17-JUL-2003  
 Endogenous nucleic fragment associated with an autoimmune disease,  
 marking method and reagent.

BD267487  
 BD267487.1 GI:33077255

JP 2002534980-A/28.  
 Homo sapiens (human)

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

1 (bases 1 to 7582)

Baccala,G.P., Mallet,F. and Voisset,C.

Endogenous nucleic fragment associated with an autoimmune disease,  
 marking method and reagent

Patent: JP 2002534980-A 28 22-OCT-2002;  
 BIO MERIEUX

OS Homo sapiens (human)  
 PN JP 2002534980-A/28  
 PD 22-OCT-2002  
 PF 21-JAN-2000 JP 2000594929  
 PR 21-JAN-1999 FR 99/00888  
 PI GLAUCIA PARANHOS BACCALA, FRANCOIS MALLET, CECILE VOISSET PC  
 C12N15/09,C07K14/15,C12Q1/02,C12Q1/68,G01N33/53, G01N33/53, PC

G01N33/564,  
 PC G01N33/566 G01N33/569 G01N33/58//C12P21/02,C12N15/00 CC  
 Endogenous nucleic fragment associated with an autoimmune CC  
 disease,  
 CC marking method and reagent  
 FH Key Location/Qualifiers  
 FT source 1..7582  
 FT Location/Qualifiers  
 FEATURES  
 source  
 1..7582  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 7582; DB 2; Length 7582;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAACAATCGGGATATAAACCAGGCAATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCC 60  
 DB 1 CAACAATCGGGATATAAACCAGGCAATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCC 60  
 QY 61 CTTCCCTTTGATGGGAGCTGTTTTCATGCTATTTCACTCTATTAATCTTCAACTGCA 120  
 DB 61 CTTCCCTTTGATGGGAGCTGTTTTCATGCTATTTCACTCTATTAATCTTCAACTGCA 120  
 QY 121 CTCCTTCGTGTCATGTTCTTACGGCTCGAGCTGAGCTTTTGCTCAACCGTCCACACTGC 180  
 DB 121 CTCCTTCGTGTCATGTTCTTACGGCTCGAGCTGAGCTTTTGCTCAACCGTCCACACTGC 180  
 QY 181 TGTTTGGCCACCGCCGACCTGCGCTGACCTCCCATCCCTCTGATCTCGCAGGGTGC 240  
 DB 181 TGTTTGGCCACCGCCGACCTGCGCTGACCTCCCATCCCTCTGATCTCGCAGGGTGC 240  
 QY 241 CGCTGTGCTCTGATCCAGCGGCGCCCAATTCGCGCTCCCAATTTGGGTAAAGGCTTCG 300  
 DB 241 CGCTGTGCTCTGATCCAGCGGCGCCCAATTCGCGCTCCCAATTTGGGTAAAGGCTTCG 300  
 QY 301 CATTTGNCCTGCACGGCTAAAGTGCCTGGGTTTGTCTAATTGAGCTGAACACTANTCACT 360  
 DB 301 CATTTGNCCTGCACGGCTAAAGTGCCTGGGTTTGTCTAATTGAGCTGAACACTANTCACT 360  
 QY 361 GGGTTCATGGTCTCTCTGTGACCCAGCGCTTCTAATAKAACTATAACACTTACCACA 420  
 DB 361 GGGTTCATGGTCTCTCTGTGACCCAGCGCTTCTAATAKAACTATAACACTTACCACA 420  
 QY 421 TGGCCCAAGATTCCATTCCTTGGAAATCCGTGAGGSCAACGAACTCCAGTCAAGAAATAC 480  
 DB 421 TGGCCCAAGATTCCATTCCTTGGAAATCCGTGAGGSCAACGAACTCCAGTCAAGAAATAC 480  
 QY 481 GARGCTTGCACCATCTTGGAAAGCGGCTGTACCTCTTGGAAAGTGTTCACCACTC 540  
 DB 481 GARGCTTGCACCATCTTGGAAAGCGGCTGTACCTCTTGGAAAGTGTTCACCACTC 540  
 QY 541 TTGGGAGCTCTGTGACGAAGACCCCGGTACATTTTGGGACCAACGAACTCCAGTCAAGAAATAC 600  
 DB 541 TTGGGAGCTCTGTGACGAAGACCCCGGTACATTTTGGGACCAACGAACTCCAGTCAAGAAATAC 600  
 QY 601 MAAGTGATGGGAACCTTCCCGCAAGACAAACCGCCCTTAAGACGTATTCTGGARAAT 660  
 DB 601 MAAGTGATGGGAACCTTCCCGCAAGACAAACCGCCCTTAAGACGTATTCTGGARAAT 660  
 QY 661 TGGGAMCAATTTGACCTTCAGACACTAAGAAAGAAACGACTTATATTTCTTCGAGTGCC 720  
 DB 661 TGGGAMCAATTTGACCTTCAGACACTAAGAAAGAAACGACTTATATTTCTTCGAGTGCC 720  
 QY 721 GCCTGGCACTCTGAGGGAAGTATAAATTATAACCATCTTACAGCTAGACTCTTTTG 780  
 DB 721 GCCTGGCACTCTGAGGGAAGTATAAATTATAACCATCTTACAGCTAGACTCTTTTG 780  
 QY 781 TAGAAAAGGCAATGGAGTGAAGTGCATAAGTACAAACTTTCTTTTCAATTAAGAGACAA 840  
 DB 781 TAGAAAAGGCAATGGAGTGAAGTGCATAAGTACAAACTTTCTTTTCAATTAAGAGACAA 840

DB 781 TAGAAAAGGCAATGGAGTGAAGTGCATAAGTACAAACTTTCTTTTCAATTAAGAGACAA 840  
 QY 841 CTCACAATTATGTAATAAAGTGTGATTTATGCCCCCTACAGGAAGCCTTCAGAGTCTACCTCC 900  
 DB 841 CTCACAATTATGTAATAAAGTGTGATTTATGCCCCCTACAGGAAGCCTTCAGAGTCTACCTCC 900  
 QY 901 CTATCCCAGCATCCCCGACTCTCTTCCCAATTAAGGACCCCTTCAACCAAAATGG 960  
 DB 901 CTATCCCAGCATCCCCGACTCTCTTCCCAATTAAGGACCCCTTCAACCAAAATGG 960  
 QY 961 TCCAAAAGGAGATAGACAAAAGGCTAAACAGTGAACCAAGAGTGCCCAATATCCCAAT 1020  
 DB 961 TCCAAAAGGAGATAGACAAAAGGCTAAACAGTGAACCAAGAGTGCCCAATATCCCAAT 1020  
 QY 1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAAATTCGGCCCGCAGCAGTGCATGTGCT 1080  
 DB 1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAAATTCGGCCCGCAGCAGTGCATGTGCT 1080  
 QY 1081 TTTTCTTCCCGAGACTTAAAGCAATATAAAGCAAGCTTTAGGTAAATCTCAGATAAATCT 1140  
 DB 1081 TTTTCTTCCCGAGACTTAAAGCAATATAAAGCAAGCTTTAGGTAAATCTCAGATAAATCT 1140  
 QY 1141 GATGGCTATATTTGTTTATCAAGGGTTAGGACAAATTTCTTTGATCTGACATGGAGAT 1200  
 DB 1141 GATGGCTATATTTGTTTATCAAGGGTTAGGACAAATTTCTTTGATCTGACATGGAGAT 1200  
 QY 1201 ATATATGTCATGCTTAATCAGACACTAAACCCCAATTAAGGAGAGTGCCACATACTGC 1260  
 DB 1201 ATATATGTCATGCTTAATCAGACACTAAACCCCAATTAAGGAGAGTGCCACATACTGC 1260  
 QY 1261 AGCCTGAGGTGTTGCGCATCTCTGCTATCTCAGTCAGTCAATGATGATGATGACAA 1320  
 DB 1261 AGCCTGAGGTGTTGCGCATCTCTGCTATCTCAGTCAGTCAATGATGATGATGACAA 1320  
 QY 1321 GAAGAAAGANAATGATTTCCCAAGGCTAGGACAAATTTCCAGTCTTCCAGTCTTCCAGTCT 1380  
 DB 1321 GAAGAAAGANAATGATTTCCCAAGGCTAGGACAAATTTCCAGTCTTCCAGTCTTCCAGTCT 1380  
 QY 1381 GGGACACAGAAATCAGTAAACATGGGAGATTGGTGTGAGACATTTGCTTAATCTGTGTGC 1440  
 DB 1381 GGGACACAGAAATCAGTAAACATGGGAGATTGGTGTGAGACATTTGCTTAATCTGTGTGC 1440  
 QY 1441 TASAGGACTTAAGAAACTAGGAAAGTCTATGAAATTAATCAATCATGTCCACCAT 1500  
 DB 1441 TASAGGACTTAAGAAACTAGGAAAGTCTATGAAATTAATCAATCATGTCCACCAT 1500  
 QY 1501 ACACAGGGAAGGGAAGAAATTCCTTCTGAGAGACTTAAGAGGAGCAATTGAG 1560  
 DB 1501 ACACAGGGAAGGGAAGAAATTCCTTCTGAGAGACTTAAGAGGAGCAATTGAG 1560  
 QY 1561 GAAGGTGCTCTCTGTCACTGACTCTTCTGAAAGGCAACTAATCTTAAAGCGTAAGTT 1620  
 DB 1561 GAAGGTGCTCTCTGTCACTGACTCTTCTGAAAGGCAACTAATCTTAAAGCGTAAGTT 1620  
 QY 1621 TATCACTCAGTCAGTCAGACATTTAGAAAACCTTCAAAAAGTCTGCGTAGGCCCCGAG 1680  
 DB 1621 TATCACTCAGTCAGTCAGACATTTAGAAAACCTTCAAAAAGTCTGCGTAGGCCCCGAG 1680  
 QY 1681 CAAAACCTTAGAAACCTTATGAACTGCGTCTGCTTTTATTAATAGAGATCAGGAG 1740  
 DB 1681 CAAAACCTTAGAAACCTTATGAACTGCGTCTGCTTTTATTAATAGAGATCAGGAG 1740  
 QY 1741 GAGCAGCGGAACAGGACAAACCGGATTTAAAAAAGGCCCGCTTTAGTCAATGACCT 1800  
 DB 1741 GAGCAGCGGAACAGGACAAACCGGATTTAAAAAAGGCCCGCTTTAGTCAATGACCT 1800  
 QY 1801 CAGCAGTGCATTTGGAGGCTCTGGAAGGGAAGGAGCTGGGCAATTTGAATGCCCTAA 1860  
 DB 1801 CAGCAGTGCATTTGGAGGCTCTGGAAGGGAAGGAGCTGGGCAATTTGAATGCCCTAA 1860  
 QY 1861 TAGGCTTGTCTTCCAGTGCAGTCTACAGGACACTTTTAAAAAAGATTTGTCAGTAGAAG 1920  
 DB 1861 TAGGCTTGTCTTCCAGTGCAGTCTACAGGACACTTTTAAAAAAGATTTGTCAGTAGAAG 1920

QY 1921 TAAGCGCCCTTCTGTCATGCGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCC 1980  
DB 1921 TAAGCGCCCTTCTGTCATGCGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCC 1980  
QY 1981 CAGGGACAAGAGTCTTTTGGAGTCAGAAAGCACTAACAGATGATCCAGCAGCAGGACTG 2040  
DB 1981 CAGGGACAAGAGTCTTTTGGAGTCAGAAAGCACTAACAGATGATCCAGCAGCAGGACTG 2040  
QY 2041 AGGGTCTGGGCAAGCGCCATCCATGCCATCAACCTCAGAGCCCTGGGTATGCTT 2100  
DB 2041 AGGGTCTGGGCAAGCGCCATCCATGCCATCAACCTCAGAGCCCTGGGTATGCTT 2100  
QY 2101 GACCATTTAGGGCCAGGAAGTGTCTCTCGGACACTGTGCGGTCTTTCTTACTTACT 2160  
DB 2101 GACCATTTAGGGCCAGGAAGTGTCTCTCGGACACTGTGCGGTCTTTCTTACTTACT 2160  
QY 2161 CTTCTGTCCGGCAAACTGTCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGAGC 2220  
DB 2161 CTTCTGTCCGGCAAACTGTCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGAGC 2220  
QY 2221 GGCACTCATAGATATCTTTTCCAGCCACTAAGTTATGAACCTGAGGAGCTTTATCTTT 2280  
DB 2221 GGCACTCATAGATATCTTTTCCAGCCACTAAGTTATGAACCTGAGGAGCTTTATCTTT 2280  
QY 2281 TCACATGCTTTTCTTAATTATGCTTGAAGCCCCACTACCTTGTGTAGGAGACATTTCTA 2340  
DB 2281 TCACATGCTTTTCTTAATTATGCTTGAAGCCCCACTACCTTGTGTAGGAGACATTTCTA 2340  
QY 2341 GCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAGGAACACCCGTTTGTGTTNC 2400  
DB 2341 GCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAGGAACACCCGTTTGTGTTNC 2400  
QY 2401 CTTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAGCAATATGAGCAGC 2460  
DB 2401 CTTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAGCAATATGAGCAGC 2460  
QY 2461 CAAGAATGCCGTCTGTTCAAGTTAACTAAGGATTCACCTCTCTTCCCTACCAAA 2520  
DB 2461 CAAGAATGCCGTCTGTTCAAGTTAACTAAGGATTCACCTCTCTTCCCTACCAAA 2520  
QY 2521 GGCAATACCCCTCAGACCCAGGCCCAACAGGATTCCAAAGATTGTTAAGCACTTA 2580  
DB 2521 GGCAATACCCCTCAGACCCAGGCCCAACAGGATTCCAAAGATTGTTAAGCACTTA 2580  
QY 2581 AAGCCCAAGCTTTAGTAAACCAATGATTAACCTCCCTGCAATATCCGTAGTGATGAG 2640  
DB 2581 AAGCCCAAGCTTTAGTAAACCAATGATTAACCTCCCTGCAATATCCGTAGTGATGAG 2640  
QY 2641 GAGGCAAGAAACCAAGTGAAGTGGAGGTTAGTGAAGATCTCAGGATTAATCAATGG 2700  
DB 2641 GAGGCAAGAAACCAAGTGAAGTGGAGGTTAGTGAAGATCTCAGGATTAATCAATGG 2700  
QY 2701 AGGCCGTGCTTTTATACCAAGCTGACCTAGCCCTTATCTGTGMYTTCCTCAATAC 2760  
DB 2701 AGGCCGTGCTTTTATACCAAGCTGACCTAGCCCTTATCTGTGMYTTCCTCAATAC 2760  
QY 2761 CAGAGGAAGCAGAGTGGTTTACASTCTCGACCTTGAAGATGCTTCTTCTGATCCCTG 2820  
DB 2761 CAGAGGAAGCAGAGTGGTTTACASTCTCGACCTTGAAGATGCTTCTTCTGATCCCTG 2820  
QY 2821 TACATCCTGACTCTCAATTTCTGTTTGGCTTTTGAAGATCTTCAAAACCCCAATCTCAAC 2880  
DB 2821 TACATCCTGACTCTCAATTTCTGTTTGGCTTTTGAAGATCTTCAAAACCCCAATCTCAAC 2880  
QY 2881 TCACCTGGACTTTTACCCCAAGGTTCAAGGATAGYCCCATCTATTTTGGCAGGCAT 2940  
DB 2881 TCACCTGGACTTTTACCCCAAGGTTCAAGGATAGYCCCATCTATTTTGGCAGGCAT 2940  
QY 2941 TAGCCCAAGACTTGAAGYCATYMTACATCTGACACTCTTCTCTGTAGTGGATG 3000  
DB 2941 TAGCCCAAGACTTGAAGYCATYMTACATCTGACACTCTTCTCTGTAGTGGATG 3000

QY 3001 ATTTACTTTTRGCGCYCCYRTTCAGAAACCTTGTGTCATCAAGCCACCCAGCRCTCTTMA 3060  
DB 3001 ATTTACTTTTRGCGCYCCYRTTCAGAAACCTTGTGTCATCAAGCCACCCAGCRCTCTTMA 3060  
QY 3061 ATTTCTCTCGYACCTGTGCTACAWGGTTTCCAAACSAARAGCTCARCTCTGCTCACAGC 3120  
DB 3061 ATTTCTCTCGYACCTGTGCTACAWGGTTTCCAAACSAARAGCTCARCTCTGCTCACAGC 3120  
QY 3121 AGGTTAAATACCTTAGGRCCTAARATTTATCCAAAGGCAACARGGCCCTCAGTGAGGAAYRYA 3180  
DB 3121 AGGTTAAATACCTTAGGRCCTAARATTTATCCAAAGGCAACARGGCCCTCAGTGAGGAAYRYA 3180  
QY 3181 TCCAGCCTATATCTGCTTATCTCATCYCAAAACCTTAAGCAACTAAGGRRTTCTTTCG 3240  
DB 3181 TCCAGCCTATATCTGCTTATCTCATCYCAAAACCTTAAGCAACTAAGGRRTTCTTTCG 3240  
QY 3241 GCRTAAYAGGYTTCTGCCGAATATGATTTCCCGAGTGGGCAAAATAGCCAGGYCATTA 3300  
DB 3241 GCRTAAYAGGYTTCTGCCGAATATGATTTCCCGAGTGGGCAAAATAGCCAGGYCATTA 3300  
QY 3301 WATACASTAATTAAGGAACTCAGAAAGCAATACCCATTTTARTAGATGGAYAMCTGAA 3360  
DB 3301 WATACASTAATTAAGGAACTCAGAAAGCAATACCCATTTTARTAGATGGAYAMCTGAA 3360  
QY 3361 GYMRAAGTGGCTTTCCAGGCCCTTAAGAGGCCCTTAACCCAGHCCCAGTGTTAAGYT 3420  
DB 3361 GYMRAAGTGGCTTTCCAGGCCCTTAAGAGGCCCTTAACCCAGHCCCAGTGTTAAGYT 3420  
QY 3421 TGCCAAACRGGGCAAGACTTTTSTTVATAVRTCAAGAAAAACAGRAAYAGCTCTRGA 3480  
DB 3421 TGCCAAACRGGGCAAGACTTTTSTTVATAVRTCAAGAAAAACAGRAAYAGCTCTRGA 3480  
QY 3481 GTCTTTACACAGRTCCRAGGAYGAGCTTGCACCCYRTGGCRVACCTGASTAAGGAAAYT 3540  
DB 3481 GTCTTTACACAGRTCCRAGGAYGAGCTTGCACCCYRTGGCRVACCTGASTAAGGAAAYT 3540  
QY 3541 GATGTAGTGGCAAGGGTTGRCYCTATTTGTTTAYGGTAGTGGTGGCAGTAGCAGTYKTA 3600  
DB 3541 GATGTAGTGGCAAGGGTTGRCYCTATTTGTTTAYGGTAGTGGTGGCAGTAGCAGTYKTA 3600  
QY 3601 GTATCTGAAGCAGTTTAAATATACAGGAGAGATCTTACTGTGGACATCTCATGAK 3660  
DB 3601 GTATCTGAAGCAGTTTAAATATACAGGAGAGATCTTACTGTGGACATCTCATGAK 3660  
QY 3661 GTGAAYRGCACTCACTGCTTAAAGGAGACTTGTGGCTGTGCAACACCTGTTTACTTAA 3720  
DB 3661 GTGAAYRGCACTCACTGCTTAAAGGAGACTTGTGGCTGTGCAACACCTGTTTACTTAA 3720  
QY 3721 TRTCAAGCTCTTATTACTTGAARGGCCAGTGTGCRACCTGTGCACTTGTCAACTCTTAAC 3780  
DB 3721 TRTCAAGCTCTTATTACTTGAARGGCCAGTGTGCRACCTGTGCACTTGTCAACTCTTAAC 3780  
QY 3781 CCAGYCNCACTTCTTCCAGCAATGAAGAAAGATARAAYATACTGTCAACAAATAAT 3840  
DB 3781 CCAGYCNCACTTCTTCCAGCAATGAAGAAAGATARAAYATACTGTCAACAAATAAT 3840  
QY 3841 TCTCAAACTATGCTCCAGTCCGAGGACCTTGTAGAGTTTCCYTTGACCTGATCCYACCTT 3900  
DB 3841 TCTCAAACTATGCTCCAGTCCGAGGACCTTGTAGAGTTTCCYTTGACCTGATCCYACCTT 3900  
QY 3901 CAACTTGTATCTGATGGAAGTTCTTTGTAGAAAAGGACTTCCGAAAGYGGGTATGC 3960  
DB 3901 CAACTTGTATCTGATGGAAGTTCTTTGTAGAAAAGGACTTCCGAAAGYGGGTATGC 3960  
QY 3961 AGTGTGCTAGTGAATTAATTTGAAGTAATCCCTCTCACTCAGGAACTAGTGTCTYA 4020  
DB 3961 AGTGTGCTAGTGAATTAATTTGAAGTAATCCCTCTCACTCAGGAACTAGTGTCTYA 4020  
QY 4021 GCTCCGAACTTAATAGCCYCAVTKGGCACTAGAAATAGGAGAGGAAAGGGYAAA 4080  
DB 4021 GCTCCGAACTTAATAGCCYCAVTKGGCACTAGAAATAGGAGAGGAAAGGGYAAA 4080  
QY 4081 TATATATACAGACTCTRARTATGCTYACCTAGTCTNTCCATGCCCATGMRCAATATGSAR 4140



4081 TATATATACAGACTCTTARTATGCTVACCTAGTCTNTCCATGCCATGMRGCAATATGSAR 4140  
 4141 AGAAGGGAATTCCTAACTTCGAGGAAACACCTATCAMAACATCAGGAAGCCATTAGGAR 4200  
 4141 AGAAGGGAATTCCTAACTTCGAGGAAACACCTATCAMAACATCAGGAAGCCATTAGGAR 4200  
 4201 ATTATTAYTGGCGWTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCTGGGGTCATCA 4260  
 4201 ATTATTAYTGGCGWTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCTGGGGTCATCA 4260  
 4261 NAAAGGAAAGRAAGGGAATASAGGAAATGCGCAAGCAKATATTGAAGCMAAAGAGC 4320  
 4261 NAAAGGAAAGRAAGGGAATASAGGAAATGCGCAAGCAKATATTGAAGCMAAAGAGC 4320  
 4321 TGCAGGCAAGACCTCCATTAAGAAATGCTTATTAACCTCCCTTAGTATAGGGTAATCC 4380  
 4321 TGCAGGCAAGACCTCCATTAAGAAATGCTTATTAACCTCCCTTAGTATAGGGTAATCC 4380  
 4381 CTTCCGGGAAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGAAACCTCAGGAG 4440  
 4381 CTTCCGGGAAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGAAACCTCAGGAG 4440  
 4441 CAGTTTCTCCCTCGGACGGTTAGCCACTGAGAGGAAATACCTTTGCCCTGCAAC 4500  
 4441 CAGTTTCTCCCTCGGACGGTTAGCCACTGAGAGGAAATACCTTTGCCCTGCAAC 4500  
 4501 TATCCAATGGAAATTAATTAACCTTCATCAAACTTTCTAGGATCGATAGCAC 4560  
 4501 TATCCAATGGAAATTAATTAACCTTCATCAAACTTTCTAGGATCGATAGCAC 4560  
 4561 CCATCARATGSCCAATCATTAATTTACTGACGAGCCCTTTTCAAACTATCAAGCARAT 4620  
 4561 CCATCARATGSCCAATCATTAATTTACTGACGAGCCCTTTTCAAACTATCAAGCARAT 4620  
 4621 AKTCAGGCGCTGTAATGTGCAARAATAATCCCTGCTATCGCAAGCTCCCTTC 4680  
 4621 AKTCAGGCGCTGTAATGTGCAARAATAATCCCTGCTATCGCAAGCTCCCTTC 4680  
 4681 AGGARAACAAACAGGCGATTTACCTGRARAARACTGGCAACTGATTTTACCACAG 4740  
 4681 AGGARAACAAACAGGCGATTTACCTGRARAARACTGGCAACTGATTTTACCACAG 4740  
 4741 CCCAACTCAGGATTTTACGATCTACTAGTCTGGGTARATCTTACGGGTTGGCA 4800  
 4741 CCCAACTCAGGATTTTACGATCTACTAGTCTGGGTARATCTTACGGGTTGGCA 4800  
 4801 RAGGCGCTCCCTGTAGGACAGAAAGGCCAAGGTAATAAAGGCACTAGTTTATGAA 4860  
 4801 RAGGCGCTCCCTGTAGGACAGAAAGGCCAAGGTAATAAAGGCACTAGTTTATGAA 4860  
 4861 ATAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTCTTTCCAG 4920  
 4861 ATAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTCTTTCCAG 4920  
 4921 GCCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGGCC 4980  
 4921 GCCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGGCC 4980  
 4981 TGAAGGCCACAGTCTCAGGAAAGTTCGAGAAATGAATGAAYACTCAAGGACATCTA 5040  
 4981 TGAAGGCCACAGTCTCAGGAAAGTTCGAGAAATGAATGAAYACTCAAGGACATCTA 5040  
 5041 AAAAGCAAAACCCAGGAAACCCACCTCACATGCGCTGYTCTGTGTGCTATAGCCTTAAAA 5100  
 5041 AAAAGCAAAACCCAGGAAACCCACCTCACATGCGCTGYTCTGTGTGCTATAGCCTTAAAA 5100  
 5101 AGAATCTGAACTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAG 5160  
 5101 AGAATCTGAACTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAG 5160  
 5161 CCCTTCATACCAATGACCTTTGTGCTTGACCCCAAGCAGGCCAATCTAGTTGACAGATCA 5220

5161 CCCTTCATAACCAATGACCTTGTGCTTGACCCAAAGACAGCCAACTTAGTTGACAGATCA 5220  
 5221 CCTCCTTAGCCAAATATCAACAAGTTCTTAAACAATTACAAGGAACCTTATCCCTCAGAG 5280  
 5221 CCTCCTTAGCCAAATATCAACAAGTTCTTAAACAATTACAAGGAACCTTATCCCTCAGAG 5280  
 5281 AGGAAAAAGAACTATTTCCACCCWMTGTCATCTGATAGTCAAGTCCCTTCTCTTAAT 5340  
 5281 AGGAAAAAGAACTATTTCCACCCWMTGTCATCTGATAGTCAAGTCCCTTCTCTTAAT 5340  
 5341 CCCATCCTTAGATACATCTTCCACCCWMTGTCATCTGATAGTCAAGTCCCTTCTCTTAAT 5400  
 5341 CCCATCCTTAGATACATCTTCCACCCWMTGTCATCTGATAGTCAAGTCCCTTCTCTTAAT 5400  
 5401 CGGTTAAGTGGCTGGAGTGGAGTCTTGATACATCACACTTGAGTCAAACTCTGGATAC 5460  
 5401 CGGTTAAGTGGCTGGAGTGGAGTCTTGATACATCACACTTGAGTCAAACTCTGGATAC 5460  
 5461 TGCCAAAGGAACCTGAAATCCAGGAGCAACGCTAGCTATTTCTGTGAAACCTCTAGAGG 5520  
 5461 TGCCAAAGGAACCTGAAATCCAGGAGCAACGCTAGCTATTTCTGTGAAACCTCTAGAGG 5520  
 5521 ATTTGGCGCTCTCTTCAACAACAACAGGAGGAAAGTAATAAATCAATAATCCGCC 5580  
 5521 ATTTGGCGCTCTCTTCAACAACAACAGGAGGAAAGTAATAAATCAATAATCCGCC 5580  
 5581 ATGSSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCTSTTCACTCTCACT 5640  
 5581 ATGSSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCTSTTCACTCTCACT 5640  
 5641 GCACCCCTCCATGCGCTGTATGACCACTAGTCTCCCTYACCMAGAGTTTCTATGGAGA 5700  
 5641 GCACCCCTCCATGCGCTGTATGACCACTAGTCTCCCTYACCMAGAGTTTCTATGGAGA 5700  
 5701 ATGACAGCTCCCGGAAATATTGATGCCCATCGTATAGAGTCTTTTAAAGGAAACCCCC 5760  
 5701 ATGACAGCTCCCGGAAATATTGATGCCCATCGTATAGAGTCTTTTAAAGGAAACCCCC 5760  
 5761 ACCTTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTGCCCCCTTTTGCATG 5820  
 5761 ACCTTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTGCCCCCTTTTGCATG 5820  
 5821 CATGCAAACTACTCTATTGACAGGAAAAATGATTAATCTTAGTTGCTCTGAGAGACTT 5880  
 5821 CATGCAAACTACTCTATTGACAGGAAAAATGATTAATCTTAGTTGCTCTGAGAGACTT 5880  
 5881 GGAGTCACTGCTGTGGACTTACTTCAACCAACTGGTATGCTGTATGATGGGGGTGGAGTT 5940  
 5881 GGAGTCACTGCTGTGGACTTACTTCAACCAACTGGTATGCTGTATGATGGGGGTGGAGTT 5940  
 5941 CAAGATCAGGCAAGGAAAAACATGTAAGAAGTAATCTCCCAACTCACCSSGGTACT 6000  
 5941 CAAGATCAGGCAAGGAAAAACATGTAAGAAGTAATCTCCCAACTCACCSSGGTACT 6000  
 6001 GGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCCGTACC 6060  
 6001 GGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCCGTACC 6060  
 6061 CATACTCGCCTGGTAAGCCTATTTAATACCACTTCACTGGGCTCCATGAGGTCTCGGCC 6120  
 6061 CATACTCGCCTGGTAAGCCTATTTAATACCACTTCACTGGGCTCCATGAGGTCTCGGCC 6120  
 6121 CAAAACCTCTACTGTTGGATATGCTTCCCTGAACTTCARGCCATATGTTTCAATC 6180  
 6121 CAAAACCTCTACTGTTGGATATGCTTCCCTGAACTTCARGCCATATGTTTCAATC 6180  
 6181 CCGTGTACCTGAACAACTGGAACAACTTCAGCAGAAAAATAACCACTTCCGTTTATGTA 6240  
 6181 CCGTGTACCTGAACAACTGGAACAACTTCAGCAGAAAAATAACCACTTCCGTTTATGTA 6240  
 6241 GGACCTCTGTGTTTCCAAATSTGAAATAACCCATACCTCAAACTCACTGTGTAAATTT 6300  
 6241 GGACCTCTGTGTTTCCAAATSTGAAATAACCCATACCTCAAACTCACTGTGTAAATTT 6300

```
QY 6301 AGCAATACTACATACACAACTCCCAATGCATCAGGTGGGTAACTCTCTCCACACAA 6360
Db 6301 AGCAATACTACATACACAACTCCCAATGCATCAGGTGGGTAACTCTCTCCACACAA 6360
QY 6361 ATAGTCTGCTACCTCCAGGAATATTTTGTCTGTGGTAACTCAGCCTATCTGTGTTTG 6420
Db 6361 ATAGTCTGCTACCTCCAGGAATATTTTGTCTGTGGTAACTCAGCCTATCTGTGTTTG 6420
QY 6421 AATGGCTCTTCAGAACTATGTGCTCTCAATCTTCTAGTGCCGCCYATGRCATCTAC 6480
Db 6421 AATGGCTCTTCAGAACTATGTGCTCTCAATCTTCTAGTGCCGCCYATGRCATCTAC 6480
QY 6481 ACTGAACAAGATTATACAGTTATGTATCAATCAAGCCCGCAACAAAGAGTACCATT 6540
Db 6481 ACTGAACAAGATTATACAGTTATGTATCAATCAAGCCCGCAACAAAGAGTACCATT 6540
QY 6541 CTTCTCTTTTGTATAGGAGCAGGAGTCTAGGTGCACTAGGTACTGGCATTTGGCGGTATC 6600
Db 6541 CTTCTCTTTTGTATAGGAGCAGGAGTCTAGGTGCACTAGGTACTGGCATTTGGCGGTATC 6600
QY 6601 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG 6660
Db 6601 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG 6660
QY 6661 GTCCGCGACTCCCTGCTCACTTGCAGAGTCACTTAACTCCCTAGCAGAGTAGTCCTT 6720
Db 6661 GTCCGCGACTCCCTGCTCACTTGCAGAGTCACTTAACTCCCTAGCAGAGTAGTCCTT 6720
QY 6721 CRAAATCGAAGACTTTAGACTTGTAAACCGCTGAGAGGGGGAACCTGTTTATTTTAA 6780
Db 6721 CRAAATCGAAGACTTTAGACTTGTAAACCGCTGAGAGGGGGAACCTGTTTATTTTAA 6780
QY 6781 GGGGAAGAATGCTGTTTATTATGTTAACTCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 6840
Db 6781 GGGGAAGAATGCTGTTTATTATGTTTAACTCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 6840
QY 6841 ATTCSAGATCGAATACAAAGTAKAGCAGAGAGCTTCGAAAACACTGGAACCTGGGCGCTC 6900
Db 6841 ATTCSAGATCGAATACAAAGTAKAGCAGAGAGCTTCGAAAACACTGGAACCTGGGCGCTC 6900
QY 6901 CTCAGCCRATGGATGCCCTGGATTCCTCCCTCTTAGGACCTCTAGCAGCTAATATTTG 6960
Db 6901 CTCAGCCRATGGATGCCCTGGATTCCTCCCTCTTAGGACCTCTAGCAGCTAATATTTG 6960
QY 6961 CTACTCCTCTTTGGACCTGTATCTTTTACCTCTTTTAACTTTCTCTCTCCAGAAATC 7020
Db 6961 CTACTCCTCTTTGGACCTGTATCTTTTACCTCTTTTAACTTTCTCTCTTTCCAGATC 7020
QY 7021 GAAGCTGTAAACTACAAATGGAGCCCAAGATGCACTCAAGACTAAGATCTACCGCAGA 7080
Db 7021 GAAGCTGTAAACTACAAATGGAGCCCAAGATGCACTCAAGACTAAGATCTACCGCAGA 7080
QY 7081 CCCCTGGACCGGCTGYTAGCCCAACGATCTGATGTTAATGACATCAAGGCAACCCCTCT 7140
Db 7081 CCCCTGGACCGGCTGYTAGCCCAACGATCTGATGTTAATGACATCAAGGCAACCCCTCT 7140
QY 7141 GAGGAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCAGCAGGAGCAGTTAGAGC 7200
Db 7141 GAGGAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCAGCAGGAGCAGTTAGAGC 7200
QY 7201 GGTSGTGGCCAACTCCCAACAGCACTTAGGTTTCTGTTGAGATGGGGACTGAGA 7260
Db 7201 GGTSGTGGCCAACTCCCAACAGCACTTAGGTTTCTGTTGAGATGGGGACTGAGA 7260
QY 7261 GACAGACTAGCTGGAATTTCTAGGCTGATTAAGAAATCCYTAAGCCCTAGGTGGGAAGGTG 7320
Db 7261 GACAGACTAGCTGGAATTTCTAGGCTGATTAAGAAATCCYTAAGCCCTAGGTGGGAAGGTG 7320
QY 7321 ACCACATCCACCTTTTAAACACGGGGCTGCACTTAGYTCACACCTGACCAATCAGAGAG 7380
Db 7321 ACCACATCCACCTTTTAAACACGGGGCTGCACTTAGYTCACACCTGACCAATCAGAGAG 7380
```

```
QY 7381 CTCACATAAATGCTAATAGGCAAGACAGGAGGTAAAGAAATAGCAATCATYTTATTGC 7440
Db 7381 CTCACATAAATGCTAATAGGCAAGACAGGAGGTAAAGAAATAGCAATCATYTTATTGC 7440
QY 7441 MTGAGAGCAGCAGCAGGAGGCAATGATCGGATATAAACCACCAAGTTCGAGCCGGCAA 7500
Db 7441 MTGAGAGCAGCAGCAGGAGGCAATGATCGGATATAAACCACCAAGTTCGAGCCGGCAA 7500
QY 7501 CGGCAACCCCTTTGGGTCCCTTCTTGTATGGGAGCTCTGTTTTCATGCTATTTCAC 7560
Db 7501 CGGCAACCCCTTTGGGTCCCTTCTTGTATGGGAGCTCTGTTTTCATGCTATTTCAC 7560
QY 7561 TCTATTAAATCTTGCARCTGCR 7582
Db 7561 TCTATTAAATCTTGCARCTGCR 7582

RESULT 3
LOCUS AX000966 7582 bp DNA linear PAT 10-MAR-2000
DEFINITION Sequence 11 from Patent WO9902696.
ACCESSION AX000966
VERSION AX000966.1 GI:7241208
KEYWORDS .
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 7582)
AUTHORS Beseme, F. and Blond, J.
TITLE ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE
DISEASES OR WITH PREGNANCY DISORDERS
JOURNAL Patent: WO 9902696-A 11 21-JAN-1999;
BIO MERIEUX (FR); BESEME FREDERIC (FR)
FEATURES
Source
1..7582
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 7582; DB 2; Length 7582;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAACAATCGGATATAAACCACCAAGGCAATTCGAGCTGGCAACAGCAGCCCTTTGGGTGC 60
Db 1 CAACAATCGGATATAAACCACCAAGGCAATTCGAGCTGGCAACAGCAGCCCTTTGGGTGC 60
QY 61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTTAAATCTTGCACACTGCA 120
Db 61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTTAAATCTTGCACACTGCA 120
QY 121 CTCCTTCTGCTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180
Db 121 CTCCTTCTGCTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC 180
QY 181 TGTTTGGCCACCAACCCGACANACCTGCGCTGACTCCCATCCCTGATCTCTGAGGCTGC 240
Db 181 TGTTTGGCCACCAACCCGACANACCTGCGCTGACTCCCATCCCTGATCTCTGAGGCTGC 240
QY 241 CGCTGTGCTCTGTATCCAGCGARGCGCCCATTCGCGCTCCCAATTTGGGCTAAAGGCTTC 300
Db 241 CGCTGTGCTCTGTATCCAGCGARGCGCCCATTCGCGCTCCCAATTTGGGCTAAAGGCTTC 300
QY 301 CATTTGTCCTGACCGGCTAAGTCCCTGGGTTTGTCTTAATGAGCTGAACACTANTCACT 360
Db 301 CATTTGTCCTGACCGGCTAAGTCCCTGGGTTTGTCTTAATGAGCTGAACACTANTCACT 360
QY 361 GGGTTCCTCATGTTCTTCTGTGACCCAGCGCTCTTAATAAACAATATACACTTACCACA 420
Db 361 GGGTTCCTCATGTTCTTCTGTGACCCAGCGCTCTTAATAAACAATATACACTTACCACA 420
QY 421 TGGCCCAAGATTCCATTCTTGGAAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAAATAC 480
```



|||||  
421 TGGCCCAAGATTCCTTCTGGAAATCGTGAGGSCAAGAACTCCAGGTCAAGAAATAC 480  
QY  
481 GARGCTTGCCACCATCTTTGGAGCGGCTGTACCTCTTTGGAAGTGGTTCACCAACCATC 540  
Db  
481 GARGCTTGCCACCATCTTTGGAGCGGCTGTACCTCTTTGGAAGTGGTTCACCAACCATC 540  
QY  
541 TTGGGAGCTCTGTGAGCAAGGACCCCGGTFACATTTTGGGBCAACMSRACGGACATCC 600  
Db  
541 TTGGGAGCTCTGTGAGCAAGGACCCCGGTFACATTTTGGGBCAACMSRACGGACATCC 600  
QY  
601 MAAGTGATGGGAAAGCTTCCCGCAAGACAAAAAGCCCTTAAGACGTATTCTGGAAT 660  
Db  
601 MAAGTGATGGGAAAGCTTCCCGCAAGACAAAAAGCCCTTAAGACGTATTCTGGAAT 660  
QY  
661 TGGGAMCAATTTGACCCCTCAGACACTAAGAAAGAAACGACTTATATTCTTGTGAGTGCC 720  
Db  
661 TGGGAMCAATTTGACCCCTCAGACACTAAGAAAGAAACGACTTATATTCTTGTGAGTGCC 720  
QY  
721 GCCTGGCACTCTGAGGGAAGTATAAATTATAACCACTCTTACAGCTAGACYTCTTTTG 780  
Db  
721 GCCTGGCACTCTGAGGGAAGTATAAATTATAACCACTCTTACAGCTAGACYTCTTTTG 780  
QY  
781 TAGAAAAAGCAATGGAGTGAAGTGCCATAAGCTACAAACTTTTCTTTTCAATTAAAGACAA 840  
Db  
781 TAGAAAAAGCAATGGAGTGAAGTGCCATAAGCTACAAACTTTTCTTTTCAATTAAAGACAA 840  
QY  
841 CTCACAATTTATGAAAAAGTGTATTTATGSCCTACAGGAAGCTTTCAGAGTCTACCTCC 900  
Db  
841 CTCACAATTTATGAAAAAGTGTATTTATGSCCTACAGGAAGCTTTCAGAGTCTACCTCC 900  
QY  
901 CTATCCAGCATCCCGACTCCTTCCGCMYTAATAAGGACCCCTTCACCCAAATGG 960  
Db  
901 CTATCCAGCATCCCGACTCCTTCCGCMYTAATAAGGACCCCTTCACCCAAATGG 960  
QY  
961 TCCAAAAGGAGATAGACAAAAGGTAAACAGTGAACCAAGAGTGCCAAATATTCGCCAAT 1020  
Db  
961 TCCAAAAGGAGATAGACAAAAGGTAAACAGTGAACCAAGAGTGCCAAATATTCGCCAAT 1020  
QY  
1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAGAAATTTGGCCCAAGCAGAGTGCATGCT 1080  
Db  
1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAGAAATTTGGCCCAAGCAGAGTGCATGCT 1080  
QY  
1081 TTTTCTCCAGACTTAAGCAATAAAAAAGAGTAAATTTCTCAGATTAATCT 1140  
Db  
1081 TTTTCTCCAGACTTAAGCAATAAAAAAGAGTAAATTTCTCAGATTAATCT 1140  
QY  
1141 GATGGCTATATTGRTGTTTTCAAGGGTTAGGACAAATCTTTGATCTGACATGGAGAGAT 1200  
Db  
1141 GATGGCTATATTGRTGTTTTCAAGGGTTAGGACAAATCTTTGATCTGACATGGAGAGAT 1200  
QY  
1201 ATATATGTCACCTGCTAAATCAGACACTAACCCCAATGAGAGAAGTGCCACCATACTGC 1260  
Db  
1201 ATATATGTCACCTGCTAAATCAGACACTAACCCCAATGAGAGAAGTGCCACCATACTGC 1260  
QY  
1261 AGCCTGAGGTTTGGCGATCTCTGGTATCTCAGTCAAGTCAATGGATANGATGACAACA 1320  
Db  
1261 AGCCTGAGGTTTGGCGATCTCTGGTATCTCAGTCAAGTCAATGGATANGATGACAACA 1320  
QY  
1321 GAAGGAAGANAATGATTTCCCAAGCAGCAGCAGGTTCCAGTCTASACCTCATTTG 1380  
Db  
1321 GAAGGAAGANAATGATTTCCCAAGCAGCAGCAGGTTCCAGTCTASACCTCATTTG 1380  
QY  
1381 GGGACACAGAAATCAGTAACATGGGAGATTGGTGTCTGCAGACATTTTGTGTC 1440  
Db  
1381 GGGACACAGAAATCAGTAACATGGGAGATTGGTGTCTGCAGACATTTTGTGTC 1440  
QY  
1441 TASAAGGACTAAGGAAAACTASGAAGAAATCTAYGAATTAATCAATGATGTCCACCAT 1500  
Db  
1441 TASAAGGACTAAGGAAAACTASGAAGAAATCTAYGAATTAATCAATGATGTCCACCAT 1500  
QY  
1501 ACACAGGGGAAGGGAAGAAAACTCTACTGCTTTCTGGAGAGACTAAGGGAGGCATTGAG 1560

Db  
1501 ACACAGGGGAAGGGAAGAAAACTCTACTGCTTTCTGGAGAGACTAAGGGAGGCATTGAG 1560  
QY  
1561 GAAGCGTGCCCTCTCTGTCACTGACTCTTTCTGAAGGCCAACTAATCTTTAAAGCGTAAGTT 1620  
Db  
1561 GAAGCGTGCCCTCTCTGTCACTGACTCTTTCTGAAGGCCAACTAATCTTTAAAGCGTAAGTT 1620  
QY  
1621 TATCACTCAGTCAGTCGAGACATTTAGAAAAAACTTCAAAAAGTCTGCGTAGGCCCGGAG 1680  
Db  
1621 TATCACTCAGTCAGTCGAGACATTTAGAAAAAACTTCAAAAAGTCTGCGTAGGCCCGGAG 1680  
QY  
1681 CAAAACTTAGAAACCTTATTGAACCTTGGCAACTTGGCACTTGGTATTAATAAGATCAGAG 1740  
Db  
1681 CAAAACTTAGAAACCTTATTGAACCTTGGCAACTTGGCACTTGGTATTAATAAGATCAGAG 1740  
QY  
1741 GAGCAGCGGGAACAGGACAAACGSGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCTT 1800  
Db  
1741 GAGCAGCGGGAACAGGACAAACGSGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCTT 1800  
QY  
1801 CAGGCAAGTGGACTTTTGGAGGCTCTGGAAAAAGGAAAAAGCTGGGCAAAATTTGAATGCTTAA 1860  
Db  
1801 CAGGCAAGTGGACTTTTGGAGGCTCTGGAAAAAGGAAAAAGCTGGGCAAAATTTGAATGCTTAA 1860  
QY  
1861 TAGGCTTTGCTTCCAGTGCAGTCTCAAGGACACTTTTAAAAAAGATTGTCCAAGTAGAAG 1920  
Db  
1861 TAGGCTTTGCTTCCAGTGCAGTCTCAAGGACACTTTTAAAAAAGATTGTCCAAGTAGAAG 1920  
QY  
1921 TAAGCGGCCCTTGTCTCATGCCCCCTTATTTTCAAGGGAATCACTGGAAAGGCCACTTGCCC 1980  
Db  
1921 TAAGCGGCCCTTGTCTCATGCCCCCTTATTTTCAAGGGAATCACTGGAAAGGCCACTTGCCC 1980  
QY  
1981 CAGGGGACAAAGGTCTTTTGTAGTCAGAGGCCACTAAACCAGATGATCCAGCAGCAGACTG 2040  
Db  
1981 CAGGGGACAAAGGTCTTTTGTAGTCAGAGGCCACTAAACCAGATGATCCAGCAGCAGACTG 2040  
QY  
2041 AGGGTGCCCTGGGGCAAGGCCATCCCATGCCATCACTTCAAGAGCCCTTGGGTATGCTT 2100  
Db  
2041 AGGGTGCCCTGGGGCAAGGCCATCCCATGCCATCACTTCAAGAGCCCTTGGGTATGCTT 2100  
QY  
2101 GACCAATTGAGGGCCAGGAAGGTGTCTCTGGACACTGGTGGGTCTTCTTAGTCTTACT 2160  
Db  
2101 GACCAATTGAGGGCCAGGAAGGTGTCTCTGGACACTGGTGGGTCTTCTTAGTCTTACT 2160  
QY  
2161 CTTCTGTCCCGACAACTGTCTCCAGATCTGTCACTATTTCTGAGGGGGTCCNTAAGAGC 2220  
Db  
2161 CTTCTGTCCCGACAACTGTCTCCAGATCTGTCACTATTTCTGAGGGGGTCCNTAAGAGC 2220  
QY  
2221 GGCAGTCACTAGATACTTTTCCCAGCCAATAAGTTATGAAGTGGGAGCTTTTATCTTT 2280  
Db  
2221 GGCAGTCACTAGATACTTTTCCCAGCCAATAAGTTATGAAGTGGGAGCTTTTATCTTT 2280  
QY  
2281 TCACATGCTTTTCTAATTTATGCTTGAAGGCCCACTACCTTGTAGGAGAGACATTTCTA 2340  
Db  
2281 TCACATGCTTTTCTAATTTATGCTTGAAGGCCCACTACCTTGTAGGAGAGACATTTCTA 2340  
QY  
2341 GCAAAAGCAGGGGCCATTATACCTGAACATAGAGAAGGAACACCCGCTTTTGTGTC 2400  
Db  
2341 GCAAAAGCAGGGGCCATTATACCTGAACATAGAGAAGGAACACCCGCTTTTGTGTC 2400  
QY  
2401 CCTGCTTGAAGGAAGAAATTAATCTCTGAAGTCTGGGCAACAGAAAGCAATATGGACGAGC 2460  
Db  
2401 CCTGCTTGAAGGAAGAAATTAATCTCTGAAGTCTGGGCAACAGAAAGCAATATGGACGAGC 2460  
QY  
2461 CAAAAGATGCCGCTCCTGTTCAAGTTAACTAAAGGATTTCCACTTCTTCTCCCTACCAA 2520  
Db  
2461 CAAAAGATGCCGCTCCTGTTCAAGTTAACTAAAGGATTTCCACTTCTTCTCCCTACCAA 2520  
QY  
2521 GGCAGTACCCCTCAGACCCAGGCCCAACAGGATTTCCAAAAGATTGTTAAGGACTTAA 2580  
Db  
2521 GGCAGTACCCCTCAGACCCAGGCCCAACAGGATTTCCAAAAGATTGTTAAGGACTTAA 2580  
QY  
2581 AAGCCCAAGGCTTAGTAAAAACCATCACTAACCTCCCTGCAAGTAAATCCGTAGTGGATTGAG 2640  
Db  
2581 AAGCCCAAGGCTTAGTAAAAACCATCACTAACCTCCCTGCAAGTAAATCCGTAGTGGATTGAG 2640

QY 2641 GAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAGAGATCTCAGGATTTATCAATGG 2700  
DB |||||  
QY 2641 GAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAGAGATCTCAGGATTTATCAATGG 2700  
DB |||||  
QY 2701 AGGCGGTGTCCTTTTATACCCAGCTGTACTAGCCCTTATCTGTGMYTTTCCCAATAC 2760  
DB |||||  
QY 2701 AGGCGGTGTCCTTTTATACCCAGCTGTACTAGCCCTTATCTGTGMYTTTCCCAATAC 2760  
DB |||||  
QY 2761 CAGAGGAAGCAGAGTGGTTTACASTCTCTGGACCTTTCAGAGTGGCTTCTTCTGCAATCCCTG 2820  
DB |||||  
QY 2761 CAGAGGAAGCAGAGTGGTTTACASTCTCTGGACCTTTCAGAGTGGCTTCTTCTGCAATCCCTG 2820  
DB |||||  
QY 2821 TACATCTGACTCTCAATCTCTGTTTGGCTTTCAGAGTACTTCAAAACCCARCATCTCAAC 2880  
DB |||||  
QY 2821 TACATCTGACTCTCAATCTCTGTTTGGCTTTCAGAGTACTTCAAAACCCARCATCTCAAC 2880  
DB |||||  
QY 2881 TCACCTGGACTRTTTTACCCCAAGGGTTTCAGGATAGTCCCATCTATTTTGGCCAGGCAT 2940  
DB |||||  
QY 2881 TCACCTGGACTRTTTTACCCCAAGGGTTTCAGGATAGTCCCATCTATTTTGGCCAGGCAT 2940  
DB |||||  
QY 2941 TAGCCCAAGACTTGAGYCARMTCACTGGAACCTTGTCTTTCRGTAKGTGGATG 3000  
DB |||||  
QY 2941 TAGCCCAAGACTTGAGYCARMTCACTGGAACCTTGTCTTTCRGTAKGTGGATG 3000  
DB |||||  
QY 3001 ATTACTTTTTRGCGCYRTTCAGAAACCTTGTGCGCATCAAGGCAACCAAGCCTCTTMA 3060  
DB |||||  
QY 3001 ATTACTTTTTRGCGCYRTTCAGAAACCTTGTGCGCATCAAGGCAACCAAGCCTCTTMA 3060  
DB |||||  
QY 3061 ATTTCCTCGCYACTGTGGCTTACAWGTTTCCAAACSBARGCTCARTCTGCTCACNGC 3120  
DB |||||  
QY 3061 ATTTCCTCGCYACTGTGGCTTACAWGTTTCCAAACSBARGCTCARTCTGCTCACNGC 3120  
DB |||||  
QY 3121 AGGTTAAATCTTAGGCTAARATTAATCCAAAGGCACCAAGGCCCTCAGTGAGGAAAYRYA 3180  
DB |||||  
QY 3121 AGGTTAAATCTTAGGCTAARATTAATCCAAAGGCACCAAGGCCCTCAGTGAGGAAAYRYA 3180  
DB |||||  
QY 3181 TCAGCCTTACTTGGCTTATCTCATCYCAAAACCTTAAAGCAACTAAGGRRTTCTTG 3240  
DB |||||  
QY 3181 TCAGCCTTACTTGGCTTATCTCATCYCAAAACCTTAAAGCAACTAAGGRRTTCTTG 3240  
DB |||||  
QY 3241 GRTAAAYAGGYTTCTCCGAATGATGATCCCGAGTGGTGGCRANAATAGCCAGGYCATTA 3300  
DB |||||  
QY 3241 GRTAAAYAGGYTTCTCCGAATGATGATCCCGAGTGGTGGCRANAATAGCCAGGYCATTA 3300  
DB |||||  
QY 3301 WATACASTAATTAAGGAACTCAGAAAGCCAATACCCATTARTTAAGATGGAAYMCTGAA 3360  
DB |||||  
QY 3301 WATACASTAATTAAGGAACTCAGAAAGCCAATACCCATTARTTAAGATGGAAYMCTGAA 3360  
DB |||||  
QY 3361 GYMRAAGTGGCTTTCCAGGCCCTTAAAGAAAGGCCCTTAAACCCCAAGYCCCAGTGTAAGYT 3420  
DB |||||  
QY 3361 GYMRAAGTGGCTTTCCAGGCCCTTAAAGAAAGGCCCTTAAACCCCAAGYCCCAGTGTAAGYT 3420  
DB |||||  
QY 3421 TGCCAACRGGGCAAGACTTTTSTYATAYRTCAAGAAAMAAACAGAAAYAGCTCTRGA 3480  
DB |||||  
QY 3421 TGCCAACRGGGCAAGACTTTTSTYATAYRTCAAGAAAMAAACAGAAAYAGCTCTRGA 3480  
DB |||||  
QY 3481 GTCCCTTACAGRTCCRAGGAYGAGCTTGCACCCYRTGTCRVACCTGASTAGGAAAYT 3540  
DB |||||  
QY 3481 GTCCCTTACAGRTCCRAGGAYGAGCTTGCACCCYRTGTCRVACCTGASTAGGAAAYT 3540  
DB |||||  
QY 3541 GATGTAGTGCAAGGGTTCRCYTCATTTTAYGGTGTGGTGGCAGTGGAGYKTA 3600  
DB |||||  
QY 3541 GATGTAGTGCAAGGGTTCRCYTCATTTTAYGGTGTGGTGGCAGTGGAGYKTA 3600  
DB |||||  
QY 3601 GTATCTGAACAGTGTAAAATAATACAGGGGAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
DB |||||  
QY 3601 GTATCTGAACAGTGTAAAATAATACAGGGGAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
DB |||||  
QY 3661 GTGAAYRGCACTACTGCTTAAAGGAGCTTGTGGCTGTGACAACTGTTTACTTAAA 3720  
DB |||||  
QY 3661 GTGAAYRGCACTACTGCTTAAAGGAGCTTGTGGCTGTGACAACTGTTTACTTAAA 3720  
DB |||||

QY 3721 TRTCAGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACTTTGTGCAACTCTTAAC 3780  
DB |||||  
QY 3721 TRTCAGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACTTTGTGCAACTCTTAAC 3780  
DB |||||  
QY 3781 CCAGYCNCAATTTCTTCCAGACAAATGAAGAAAGATARAAYATAACTGTCAACAARTAAAT 3840  
DB |||||  
QY 3781 CCAGYCNCAATTTCTTCCAGACAAATGAAGAAAGATARAAYATAACTGTCAACAARTAAAT 3840  
DB |||||  
QY 3841 TCTCAAAACCTATGCGCACTCGAGGGACCTTCTTGTAGARGTTCCYTTGACTGATCCYACCTT 3900  
DB |||||  
QY 3841 TCTCAAAACCTATGCGCACTCGAGGGACCTTCTTGTAGARGTTCCYTTGACTGATCCYACCTT 3900  
DB |||||  
QY 3901 CAACCTTGATATCTGATGGAAGTTCTTCTGTAGAAAGACACTTCGAAAGYGGGTATGC 3960  
DB |||||  
QY 3901 CAACCTTGATATCTGATGGAAGTTCTTCTGTAGAAAGACACTTCGAAAGYGGGTATGC 3960  
DB |||||  
QY 3961 AGTGGTCACTGATTAATGGAATATTTGAAAGTAAATCCCTCACTCCAGGAACCTAGTCTTAA 4020  
DB |||||  
QY 3961 AGTGGTCACTGATTAATGGAATATTTGAAAGTAAATCCCTCACTCCAGGAACCTAGTCTTAA 4020  
DB |||||  
QY 4021 GCTRGCAAGAACTTAATAGCCYTCAYTKGGCACCTAGAAATTTAGGAGAAAGAAAGGAAAA 4080  
DB |||||  
QY 4021 GCTRGCAAGAACTTAATAGCCYTCAYTKGGCACCTAGAAATTTAGGAGAAAGAAAGGAAAA 4080  
DB |||||  
QY 4081 TATATATACAGACTCTTARTATGCTTACCTAGTCTGATGCGCATGMRGCAATATGSA 4140  
DB |||||  
QY 4081 TATATATACAGACTCTTARTATGCTTACCTAGTCTGATGCGCATGMRGCAATATGSA 4140  
DB |||||  
QY 4141 AGAAGGGAAATTCCTAACTTCYAGRGACACCTATCAMACATCAGGAAGCAATATAGGAR 4200  
DB |||||  
QY 4141 AGAAGGGAAATTCCTAACTTCYAGRGACACCTATCAMACATCAGGAAGCAATATAGGAR 4200  
DB |||||  
QY 4201 ATTATTAYTGGCWGTACAGAAACCTTARAGAGTGGMAGTCTTACACTGCGYGGGTCTATCA 4260  
DB |||||  
QY 4201 ATTATTAYTGGCWGTACAGAAACCTTARAGAGTGGMAGTCTTACACTGCGYGGGTCTATCA 4260  
DB |||||  
QY 4261 NAAAGAAAGRAAGGAAATASAGRGAAATGCGCAAGCAKATATTTGAAGCAAAAGAGC 4320  
DB |||||  
QY 4261 NAAAGAAAGRAAGGAAATASAGRGAAATGCGCAAGCAKATATTTGAAGCAAAAGAGC 4320  
DB |||||  
QY 4321 TCAGAGGAGGACCTCCATTAAGAAATGCTTAACTTCCCTTAGTATAGGTATATCC 4380  
DB |||||  
QY 4321 TCAGAGGAGGACCTCCATTAAGAAATGCTTAACTTCCCTTAGTATAGGTATATCC 4380  
DB |||||  
QY 4381 CTTTCGGGAAACCAAGCCCGAGTACTCAGCAGAGAAACAGAAATGGGAACTTCAAGAGG 4440  
DB |||||  
QY 4381 CTTTCGGGAAACCAAGCCCGAGTACTCAGCAGAGAAACAGAAATGGGAACTTCAAGAGG 4440  
DB |||||  
QY 4441 CAGTTTTCTCCCTCGGGACGGTTAGCCACTGAAAGGAGGAAATACTTTTGCCTGCAAC 4500  
DB |||||  
QY 4441 CAGTTTTCTCCCTCGGGACGGTTAGCCACTGAAAGGAGGAAATACTTTTGCCTGCAAC 4500  
DB |||||  
QY 4501 TATCCATGGAATTAATTAAACCCCTTCAATCAACCTTTTCACTTAGGCAATGATAGCAC 4560  
DB |||||  
QY 4501 TATCCATGGAATTAATTAAACCCCTTCAATCAACCTTTTCACTTAGGCAATGATAGCAC 4560  
DB |||||  
QY 4561 CCATCARATGGCCAAATCAATTAATTTAATGGAACAGGCCCTTTTCAAAACTATCAAGCARAT 4620  
DB |||||  
QY 4561 CCATCARATGGCCAAATCAATTAATTTAATGGAACAGGCCCTTTTCAAAACTATCAAGCARAT 4620  
DB |||||  
QY 4621 AKTCAAGGGCTGTGAATGTGCAARAAATTAATCCCTGCTTATTCGCAAGCTCTTTC 4680  
DB |||||  
QY 4621 AKTCAAGGGCTGTGAATGTGCAARAAATTAATCCCTGCTTATTCGCAAGCTCTTTC 4680  
DB |||||  
QY 4681 AGGAAACAAARAAACAGGCCATTAAGGCAATTAAGGCAATTAAGGCAATTAAGGCAATTAAGGCA 4740  
DB |||||  
QY 4681 AGGAAACAAARAAACAGGCCATTAAGGCAATTAAGGCAATTAAGGCAATTAAGGCAATTAAGGCA 4740  
DB |||||  
QY 4741 CCCAAACCTCAGGATTTTCACTTACTAGTCTGGTARATCTTCAAGGCTTGGGCA 4800  
DB |||||  
QY 4741 CCCAAACCTCAGGATTTTCACTTACTAGTCTGGTARATCTTCAAGGCTTGGGCA 4800  
DB |||||  
QY 4801 RAGGCTTCCCTCTAGGACAGAAAGGCCCAAGAGGTAATAAGGCACTAGTCTCATGAA 4860  
DB |||||

|||||  
4801 RAGGCTTCCCCTGTAGGACAGAAAAGCCAAAGAGTAAATAAGGCACCTAGTTCAATGNA 4860  
|||||  
4861 AATAATCCAGATTCGGACTTCCCAGAGCTTACAGAGTGACAAATAGGCCCTCTTTCCAG 4920  
|||||  
4861 AATAATCCAGATTCGGACTTCCCAGAGCTTACAGAGTGACAAATAGGCCCTCTTTCCAG 4920  
|||||  
4921 GCCACAGTAACCCAGGAGTATCCAGAGCTTACAGAGTGATATACGATATACACTTACACTGCGCC 4980  
|||||  
4921 GCCACAGTAACCCAGGAGTATCCAGAGCTTACAGAGCTTACAGAGTGATATACGATATACACTTACACTGCGCC 4980  
|||||  
4981 TGAAGGCCACAGTCTCTCAGGAGGTCGAGAAAATGAATGAAYACTCAAAGGACATCTA 5040  
|||||  
4981 TGAAGGCCACAGTCTCTCAGGAGGTCGAGAAAATGAATGAAYACTCAAAGGACATCTA 5040  
|||||  
5041 AAAAAAGCAACCCAGGAAACCCACCTCA CATGCCCTGYTCTGTGCTATAGCCTTAAAA 5100  
|||||  
5041 AAAAAAGCAACCCAGGAAACCCACCTCA CATGCCCTGYTCTGTGCTATAGCCTTAAAA 5100  
|||||  
5101 AGAATCTGCAACTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGG 5160  
|||||  
5101 AGAATCTGCAACTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGG 5160  
|||||  
5161 CCTTTCATACCAATGACCTTGTGCTTGACCAAGACGCAACTTAGTTGACAGATCA 5220  
|||||  
5161 CCTTTCATACCAATGACCTTGTGCTTGACCAAGACGCAACTTAGTTGACAGATCA 5220  
|||||  
5221 CCTCTTAGCCAAATATCAACAGTTCTTAAACATTACAAGAACTTATCCCTGAGAAG 5280  
|||||  
5221 CCTCTTAGCCAAATATCAACAGTTCTTAAACATTACAAGAACTTATCCCTGAGAAG 5280  
|||||  
5281 AGGAAAAGAACTATTTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCTCTAATT 5340  
|||||  
5281 AGGAAAAGAACTATTTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCTCTAATT 5340  
|||||  
5341 CCCATCCCTAGATACATCTCGGAAGGACCTTACCAGTCAATTTATVTAACCCAACTG 5400  
|||||  
5341 CCCATCCCTAGATACATCTCGGAAGGACCTTACCAGTCAATTTATVTAACCCAACTG 5400  
|||||  
5401 CGGTTAAAGTGCTGGAGTGAGTCTTGATACATACACTTTGAGTCAAAATCCTGGATAC 5460  
|||||  
5401 CGGTTAAAGTGCTGGAGTGAGTCTTGATACATACACTTTGAGTCAAAATCCTGGATAC 5460  
|||||  
5461 TGGCAAGGAACCTGAAATCCAGGAGACAAAGCTAGCTATTCTCTGTGAACCTCTAGAGG 5520  
|||||  
5461 TGGCAAGGAACCTGAAATCCAGGAGACAAAGCTAGCTATTCTCTGTGAACCTCTAGAGG 5520  
|||||  
5521 ATTTGGCCTGCTCTTCAAACCAACACAGGAGAAAGTAACTAAAATCATAAATCCCCC 5580  
|||||  
5521 ATTTGGCCTGCTCTTCAAACCAACACAGGAGAAAGTAACTAAAATCATAAATCCCCC 5580  
|||||  
5581 ATGSGCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCCTSTTTCACTCTCACT 5640  
|||||  
5581 ATGSGCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCCTSTTTCACTCTCACT 5640  
|||||  
5641 GCACCCCTCCATGCGCTGTATGACAGTAGCTCCCTCYAACMAGAGTTTCTATGGAGA 5700  
|||||  
5641 GCACCCCTCCATGCGCTGTATGACAGTAGCTCCCTCYAACMAGAGTTTCTATGGAGA 5700  
|||||  
5701 ATGCAGGTCCTGGAAATATGATGCCCATCGTATAGAGTCTTTSTAAGGGAAACCCCC 5760  
|||||  
5701 ATGCAGGTCCTGGAAATATGATGCCCATCGTATAGAGTCTTTSTAAGGGAAACCCCC 5760  
|||||  
5761 ACCTTCACTGCCACACCCATATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820  
|||||  
5761 ACCTTCACTGCCACACCCATATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820  
|||||  
5821 CATGAAATCTCATTTATGGACAGGAAATGAATTAATCTAGTTGTCTGTGAGGACTT 5880  
|||||  
5821 CATGAAATCTCATTTATGGACAGGAAATGAATTAATCTAGTTGTCTGTGAGGACTT 5880  
|||||  
5881 GGAGTCACTGTCTGTGGACTTACTTCAACCAACTGGTATGCTGATGCGGGGTGGAGTT 5940  
|||||

Db 5881 GGAGTCACTGTCTGTGGACTTACTTTACCCAAACTGCTGATGCTGTAGGGGTGGAGTT 5940  
QY 5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAGTAATCTCCCACTCACTCCSGGGTACAT 6000  
|||||  
5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAGAGTAATCTCCCACTCACTCCSGGGTACAT 6000  
|||||  
6001 GGCACCTCTAGCCCTACAAAAGGACTAGATCTCTCAAAACTACATGAAAACCTTCGGTACC 6060  
|||||  
6001 GGCACCTCTAGCCCTACAAAAGGACTAGATCTCTCAAAACTACATGAAAACCTTCGGTACC 6060  
|||||  
6061 CATACTCGCTGGTAAGCCTATTTAATACACCCCTCACTCGGCTCCATGAGGTCTCGGCC 6120  
|||||  
6061 CATACTCGCTGGTAAGCCTATTTAATACACCCCTCACTCGGCTCCATGAGGTCTCGGCC 6120  
|||||  
6121 CAAAACCTCTACTGTGTGGATATGCTCCCTGTAACCTTCAGCCCATATGTTTTCAATC 6180  
|||||  
6121 CAAAACCTCTACTGTGTGGATATGCTCCCTGTAACCTTCAGCCCATATGTTTTCAATC 6180  
|||||  
6181 CCTGTACCTGAACATGGAACAATTTTCAGCAGAGAAATAAACACCACTTCCTGTTTTAGTA 6240  
|||||  
6181 CCTGTACCTGAACATGGAACAATTTTCAGCAGAGAAATAAACACCACTTCCTGTTTTAGTA 6240  
|||||  
6241 GGACCTCTGTTTCCAAATSTGGAATAACCCATACCTCAAACTTCACCTGTGTAATAATTT 6300  
|||||  
6241 GGACCTCTGTTTCCAAATSTGGAATAACCCATACCTCAAACTTCACCTGTGTAATAATTT 6300  
|||||  
6301 AGCAATATCATATACAAACCAATCCCAATGCAATCAGGTGGGTAACTCTCTCCCAACAA 6360  
|||||  
6301 AGCAATATCATATACAAACCAATCCCAATGCAATCAGGTGGGTAACTCTCTCCCAACAA 6360  
|||||  
6361 ATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCCTATCTGTTGTTTG 6420  
|||||  
6361 ATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCCTATCTGTTGTTTG 6420  
|||||  
6421 AATGCTCTTTCAGAACTCTATGTGCTTCTCATTTTAGTGCCCCCTCAATGRCATCTAC 6480  
|||||  
6421 AATGCTCTTTCAGAACTCTATGTGCTTCTCATTTTAGTGCCCCCTCAATGRCATCTAC 6480  
|||||  
6481 ACTGAAAGAATTTATACAGTTATGTCATATCTAAGCCCGCCGCAACAAAGAGTACCCATT 6540  
|||||  
6481 ACTGAAAGAATTTATACAGTTATGTCATATCTAAGCCCGCCGCAACAAAGAGTACCCATT 6540  
|||||  
6541 CTTCCCTTTGTTATAGGAGCAGGAGTGTAGTGCACTAGGTACTTGGCATTGCGGTATC 6600  
|||||  
6541 CTTCCCTTTGTTATAGGAGCAGGAGTGTAGTGCACTAGGTACTTGGCATTGCGGTATC 6600  
|||||  
6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660  
|||||  
6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660  
|||||  
6661 GTGCGGAGTCCCTGGTCACTTGCAGAGTCAACTTAATCTCCCTAGCAGCAGTAGTCTTT 6720  
|||||  
6661 GTGCGGAGTCCCTGGTCACTTGCAGAGTCAACTTAATCTCCCTAGCAGCAGTAGTCTTT 6720  
|||||  
6721 CBAATCCGAAGAGCTTTAGACTTGTAAACCGCTGARAGAGGGGAAACCTGTTTTATTTTAA 6780  
|||||  
6721 CBAATCCGAAGAGCTTTAGACTTGTAAACCGCTGARAGAGGGGAAACCTGTTTTATTTTAA 6780  
|||||  
6781 GGGGAAGAAATGCTGTTTATTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAA 6840  
|||||  
6781 GGGGAAGAAATGCTGTTTATTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAA 6840  
|||||  
6841 ATTCAGATTCGAATCAACGTAAGAGAGTCTGAAAACACTGGGACCTTGGGCGCTC 6900  
|||||  
6841 ATTCAGATTCGAATCAACGTAAGAGAGTCTGAAAACACTGGGACCTTGGGCGCTC 6900  
|||||  
6901 CTCAGCCATGATGCTGATGCTGATGCTTCTAGGACCTCTAGCAGCTAATATTTG 6960  
|||||  
6901 CTCAGCCATGATGCTGATGCTGATGCTTCTAGGACCTCTAGCAGCTAATATTTG 6960  
|||||  
6961 CTACTCTCTTTGGACCTCTTATCTTTTACCTCTTGTAACTTTGTCTCTTCCAGAAATC 7020  
|||||  
6961 CTACTCTCTTTGGACCTCTTATCTTTTACCTCTTGTAACTTTGTCTCTTCCAGAAATC 7020  
|||||

```

QY 7021 GAAGCTGTRAACTACAAATGGAGCCCAAGATGAGTCCAAAGACTAAGATCTACCGAGA 7080
Db 7021 GAAGCTGTRAACTACAAATGGAGCCCAAGATGAGTCCAAAGACTAAGATCTACCGAGA 7080
QY 7081 CCCCTGGACCGGCTGTAGCCCAAGATCTGATGTTAAATGACATCAAAAGGACACCCCTCCT 7140
Db 7081 CCCCTGGACCGGCTGTAGCCCAAGATCTGATGTTAAATGACATCAAAAGGACACCCCTCCT 7140
QY 7141 GAGGAATCTCAGCTGCAAACTCTACTAGCCCAATTCAGAGAGAAAGAGTATAGAGC 7200
Db 7141 GAGGAATCTCAGCTGCAAACTCTACTAGCCCAATTCAGAGAGAAAGAGTATAGAGC 7200
QY 7201 GGTSGTCGGCAACCTCCCAACAGCACTTAGTCTTCTGTTGAGATGGGGAGCTGAGA 7260
Db 7201 GGTSGTCGGCAACCTCCCAACAGCACTTAGTCTTCTGTTGAGATGGGGAGCTGAGA 7260
QY 7261 GACAGGACTAGCTGGATTTCTTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTG 7320
Db 7261 GACAGGACTAGCTGGATTTCTTAGGCTGAYTAAGAATCCYTAAGCCTAGSTGGGAAGGTG 7320
QY 7321 ACCACATCCACCTTTAAACACAGGGCTTGCAACTTAGYTCAACCTGACCAATCAGAGAG 7380
Db 7321 ACCACATCCACCTTTAAACACAGGGCTTGCAACTTAGYTCAACCTGACCAATCAGAGAG 7380
QY 7381 CTCACATAAATGCTAATAGGCAAGACAGGAGTAAAGAAATAGCAATCATYTTATTC 7440
Db 7381 CTCACATAAATGCTAATAGGCAAGACAGGAGTAAAGAAATAGCAATCATYTTATTC 7440
QY 7441 MTGAGAGCAGCAGGAGGAGCAATGATCGGATATAAACCCAAAGTTCGAGCCGGCAA 7500
Db 7441 MTGAGAGCAGCAGGAGGAGCAATGATCGGATATAAACCCAAAGTTCGAGCCGGCAA 7500
QY 7501 CGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTCATGCTATTTCAC 7560
Db 7501 CGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTCATGCTATTTCAC 7560
QY 7561 TCTATTAAATCTTGCARCTGCR 7582
Db 7561 TCTATTAAATCTTGCARCTGCR 7582

```

```

RESULT 4
AX027480
LOCUS AX027480 7582 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 30 from Patent FR2788784.
ACCESSION AX027480
VERSION AX027480.1 GI:10188444
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Mallet,F., Voisset,C. and Paranhos,B.G.
AUTHORS Patient: FR 2788784-A 30 28-JUL-2000;
JOURNAL BIO MERIEUX (FR)
FEATURES
    Location/Qualifiers
        1..7582
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 7582; DB 2; Length 7582;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCTTTGGGTCC 60
Db 1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCTTTGGGTCC 60

```

```

QY 61 CTTCCCTTTGTATGGAGCTGTTTTCATGCTATTTCACTCTATTAATCTTGCAACTGCA 120
Db 61 CTTCCCTTTGTATGGAGCTGTTTTCATGCTATTTCACTCTATTAATCTTGCAACTGCA 120
QY 121 CTTCTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTCACCGTCCACCACTGC 180
Db 121 CTTCTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTCACCGTCCACCACTGC 180
QY 181 TGTTCGCAACCCGACANACCTGCGCTGACTCCCATCCCTCTGGATCTCTGCAAGGTGTC 240
Db 181 TGTTCGCAACCCGACANACCTGCGCTGACTCCCATCCCTCTGGATCTCTGCAAGGTGTC 240
QY 241 CGCTGTGCTCCTGATCCAGGARGCGCCCAATTGCGCTCCCAATTGGGCTAAAGGCTTGC 300
Db 241 CGCTGTGCTCCTGATCCAGGARGCGCCCAATTGCGCTCCCAATTGGGCTAAAGGCTTGC 300
QY 301 CATTTGTCCTGCAACGGCTTAAGTCCCTGGGTTTGTCTTAATTGAGCTGAAACACTANTCACT 360
Db 301 CATTTGTCCTGCAACGGCTTAAGTCCCTGGGTTTGTCTTAATTGAGCTGAAACACTANTCACT 360
QY 361 GGGTTCCATGGTTCCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA 420
Db 361 GGGTTCCATGGTTCCTTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA 420
QY 421 TGGCCCCAAGATTCATCTTCTTGGATCCGCTGAGGSCAAAGAACTCCAGGTCAGAGAAATAC 480
Db 421 TGGCCCCAAGATTCATCTTCTTGGATCCGCTGAGGSCAAAGAACTCCAGGTCAGAGAAATAC 480
QY 481 GARGCTTGGCACCATCTTGGAGCGGCTGCTACRTCTTGGAGTGGTTCACCAACATC 540
Db 481 GARGCTTGGCACCATCTTGGAGCGGCTGCTACRTCTTGGAGTGGTTCACCAACATC 540
QY 541 TTGGGAGCTCTGTGAGCAAGGACCCCGGCTTACATTTTGGCRACCAACGAGCATCC 600
Db 541 TTGGGAGCTCTGTGAGCAAGGACCCCGGCTTACATTTTGGCRACCAACGAGCATCC 600
QY 601 MAAGTATGGGAAAACGTTCCCGCAAGACAAAACGCGCTTAAGAGCTATTTCTGGARAAT 660
Db 601 MAAGTATGGGAAAACGTTCCCGCAAGACAAAACGCGCTTAAGAGCTATTTCTGGARAAT 660
QY 661 TGGGAMCAATTTGACCTCAGACACTAAGAAAGAAAGAGCTTATATTTCTTCGAGTGC 720
Db 661 TGGGAMCAATTTGACCTCAGACACTAAGAAAGAAAGAGCTTATATTTCTTCGAGTGC 720
QY 721 GCTTGCACCTCTGAGGGAAGTATAATTAACACCATCTTACAGCTAGACTCTTTTTC 780
Db 721 GCTTGCACCTCTGAGGGAAGTATAATTAACACCATCTTACAGCTAGACTCTTTTTC 780
QY 781 TAGAAAAGGCAATGGAGTGAAGTGCATTAAGTACAAACTTTTCTTTCATTAAGAGACAA 840
Db 781 TAGAAAAGGCAATGGAGTGAAGTGCATTAAGTACAAACTTTTCTTTCATTAAGAGACAA 840
QY 841 CTCACAAATTTATGAAAAAGTGTGATTTATGCGCTCAGAGAGGCTTCAGAGTCTACTCC 900
Db 841 CTCACAAATTTATGAAAAAGTGTGATTTATGCGCTCAGAGAGGCTTCAGAGTCTACTCC 900
QY 901 CTATCCAGCATCCCGACTCTCTTCCCAAMYTAAATGAAGACCCCTTCAACCCCAATGG 960
Db 901 CTATCCAGCATCCCGACTCTCTTCCCAAMYTAAATGAAGACCCCTTCAACCCCAATGG 960
QY 961 TCCAAAAGGAGATAGACAAAAGGTAACAGTGAACCAAGAGTGCATATTTCCCAAT 1020
Db 961 TCCAAAAGGAGATAGACAAAAGGTAACAGTGAACCAAGAGTGCATATTTCCCAAT 1020
QY 1021 TATGACCCCTCCCAAGCAGTGGAGAGAGAAATTCGGCCCAAGCAGAGTGCATGTGCT 1080
Db 1021 TATGACCCCTCCCAAGCAGTGGAGAGAGAAATTCGGCCCAAGCAGAGTGCATGTGCT 1080
QY 1081 TTTTCTTCCAGACTTAAAGCAATAAATAAAGAGTAAATTTCTCAGATAATCCCT 1140
Db 1081 TTTTCTTCCAGACTTAAAGCAATAAATAAAGAGTAAATTTCTCAGATAATCCCT 1140
QY 1141 GATGGCTATATTTGTTTACAAAGGTTAGGACAATTTCTTGTATCTGACATGGAGAT 1200

```





Qy	3361	GYMBAAGTGGCTTTTCAGGCGCCCTTAAAGAGGCCCTTAAACCCAGYCCCAGTGTAAAGYT	3420
Db	3361		
Qy	3421	TGCCAACRGGGCAAGACTTTTSTTYATAYRTACAGAAAAAAGCAAGRAAVAGCTCTRGGA	3480
Db	3421		
Qy	3481	GTCTTACACAGRTCCRAGGAGYAGCTTGGACCCYRTGGGRVACCTGASTAAGGAAAYT	3540
Db	3481		
Qy	3541	GATGTAGTGGCAAGGGTTCRCYCTATGTTTAYGGTAGTGGTGGCAGTAGCAGTYKTA	3600
Db	3541		
Qy	3601	GTATCTGAAGCAGTAAATAATAACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK	3660
Db	3601		
Qy	3661	GTGAAVYRGATACCTACTGTCTAAAGGAGACTTGTGGCTGTGACAACTCTTTAA	3720
Db	3661		
Qy	3721	TRTCAGGCTCTATTACTTGAAGGCCAGTCTGTGACTGTGCACTTGTGCAACTCTTAAC	3780
Db	3721		
Qy	3781	CCAGYCNCACTTCTCCAGACAATGAAGAAAGATAPAAATACCTGTCAACAARTAAAT	3840
Db	3781		
Qy	3841	TCTCAAACTATGCCACTCGAGGGGACCTTCTTAGAGTTCCYTTGACTGATCCYGACCTT	3900
Db	3841		
Qy	3901	CAACTTGTATCTGATGGAAGTCTTTTGTAGAAAAGGACTTCGAAAAGYGGGGTATGC	3960
Db	3901		
Qy	3961	AGTGGTCAGTGATAATAGCCYTCAATYTRGGGCACTAGAAATAGGAGAAAGAAAAGGYAAA	4020
Db	3961		
Qy	4021	GCTRGCAAGCACTAATAGCCYTCAATYTRGGGCACTAGAAATAGGAGAAAGAAAAGGYAAA	4080
Db	4021		
Qy	4081	TATATATACAGACTCTTRATATGCTTACCTAGTCTCCATGCCCATTGMRCAATATGSAR	4140
Db	4081		
Qy	4141	AGAAAGGGAATCCTTAATCTCYAGRGACACCTATCMAACATCAGGAAGCCATTAGGAR	4200
Db	4141		
Qy	4201	ATTATTAATYTGCGWGTACAGAAACCTARAGAGGTGGMAGTCTTACATGCTGCGGGTCTATCA	4260
Db	4201		
Qy	4261	NAAGGAAAGRAAAGGGAATASAGRGAAATGCCAAGCAKATATTTGAAGCMAAAAGAGC	4320
Db	4261		
Qy	4321	TGCAAGGCAAGCCCTCCATTAGAAATGCTTATTAACCTTCCCTTATGATAGGGTAATCC	4380
Db	4321		
Qy	4381	CTTCCGGGAAACCAAGCCCGAGTACTCAGCAGAGAGAAACAGAAATGGGGACCTCAGAGG	4440
Db	4381		

Qy	4441	CAGTTTTCTCCCTCGGACGGTTAGCCACTGAAGAGGAAATACTTTTGCCTGCAAC	4500
Db	4441		
Qy	4501	TATCCAATGAAATTAATTAATAAACCCTTCAATCAAACTTTTCACTTAGGCATCGATAGCAC	4560
Db	4501		
Qy	4561	CCATCARATGGCCAAATCATATTATTTACTGGACAGGCTTTTCAAACTATCAAGCARAT	4620
Db	4561		
Qy	4621	AKTCAGGCTGTGAATGTGCCARARAAATAATCCCTGCCCTTATGCGCAAGCTCTTC	4680
Db	4621		
Qy	4681	AGGARAACAARAACAGGCCATTACCTGRRARARACTGGCAACTGATTTTATCCCAAG	4740
Db	4681		
Qy	4741	CCCAAACTCAGGGATTTCAGTATCTACTAGTCTGGGTARATATCTTTCACGGTTGGGCA	4800
Db	4741		
Qy	4801	RAGGCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTATTAAGGCACTAGTTTCATGAA	4860
Db	4801		
Qy	4861	ATAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGTTTCCAG	4920
Db	4861		
Qy	4921	GCACAGTAACCCAGGGAGTATCCAGGGCTTAGGTATACGATATCACTTACACTGCGCC	4980
Db	4921		
Qy	4981	TGAAGGCCACAGTCTCTCAGGGAAGGTGAGAAAATGAATGAAAYATCAAAAGACATCTA	5040
Db	4981		
Qy	5041	AAAAAGCAAAACCCAGGAAACCCACCTCACTGCGCTGTCTGTTGCTATAGCCTTAAAA	5100
Db	5041		
Qy	5101	AGAACTCGCAACTTTTCCCAAAAGCAGCACTTAGGCCATACGAAATGCTGTATGGAAG	5160
Db	5101		
Qy	5161	CCCTTCATAACCAATGACCTTGTGCTTGACCCCAAGACAGCCAACTTAGTTGCAGACATCA	5220
Db	5161		
Qy	5221	CTCTCTTAGCCAAATATCAAAAGTCTTTAAACATTAAGGAACCTATCCCTGAGAAG	5280
Db	5221		
Qy	5281	AGGAAAAAGAACTATTTCACCCWGTGACATGATATTAGTCAAGTCCCTTCCTCTAAT	5340
Db	5281		
Qy	5341	CCCCATCCCTAGATACATCTCGGAGGACCCCTACCCAGTCACTTTATTTACCCCACTG	5400
Db	5341		
Qy	5401	CGGTTAAAGTGGTGGAGTGTCTTGATATCATCACTTGAGTCAAAATCTTGATAC	5460
Db	5401		
Qy	5461	TGCCAAAGGAACTGAAATCCAGAGACAAAGCTAGCTATTCTCTGTGAACTCTAGAGG	5520
Db	5461		
Qy	5521	ATTTGCGCTGCTCTTCAAAACCAACAGGAGGAAAGTAACTAAATCATAAATCCCCC	5580
Db	5521		

[illegible]

	Db	6601	ACAACTCTACTCAGTTTCTACTACAAACTATCTCAAGAACTAAATGGGGAATGAG3AACGG	6666
Qy	6661	GTCGCCAGCTCCCTGGTGACCTTGCAAGATCAAACTTAACCTCCCCTAGCAGCAGTAGTCCTTT	6720	
Db	6661	GTCGCCAGCTCCCTGGTGACCTTGCAAGATCAAACTTAACCTCCCCTAGCAGCAGTAGTCCTTT	6720	
Qy	6721	CRAAATCGAAGAGCTTTAGACTVGCCTAACCGCTGARAGAGGGGAACTGTTTATTATTTTTA	6780	
Db	6721	CRAAATCGAAGAGCTTTAGACTVGCCTAACCGCTGARAGAGGGGAACTGTTTATTATTTTTA	6780	
Qy	6781	GGGGAAGAAGCTGTATTATTTAATCATCCGGATCCGTCACTGAGAAGTTTRAGAA	6840	
Db	6781	GGGGAAGAAGCTGTATTATTTAATCATCCGGATCCGTCACTGAGAAGTTTRAGAA	6840	
Qy	6841	ATTCAGATCGAATAACAACGTAKAGCAGARGAGCTTCGAAAACACTGGACCCCTGGGGCCTC	6900	
Db	6841	ATTCAGATCGAATAACAACGTAKAGCAGARGAGCTTCGAAAACACTGGACCCCTGGGGCCTC	6900	
Qy	6901	CTCAGCCRATGGATGCCCTGGATTCCTCCCTTTTAGGACCTCTAGCAGCTATAATATTG	6960	
Db	6901	CTCAGCCRATGGATGCCCTGGATTCCTCCCTTTTAGGACCTCTAGCAGCTATAATATTG	6960	
Qy	6961	CTACTCCTCTTTGGACCCTGTATCTTTTACCTCCTTGTTAACTTTGTCTCTCCAGAACT	7020	
Db	6961	CTACTCCTCTTTGGACCCTGTATCTTTTACCTCCTTGTTAACTTTGTCTCTCCAGAACT	7020	
Qy	7021	GAAGCTGRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA	7080	
Db	7021	GAAGCTGRAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGA	7080	
Qy	7081	CCCCTGGACCGGCTGYTAGCCCCAGACTCTGATGTTTAATGACACTCAAAGCACCCCTCCT	7140	
Db	7081	CCCCTGGACCGGCTGYTAGCCCCAGACTCTGATGTTTAATGACACTCAAAGCACCCCTCCT	7140	
Qy	7141	GAGGAATCTCAGCTGCACAACTCTACTAGCCGCCAAATTCAGCAGGAAGCAGTTAGAGC	7200	
Db	7141	GAGGAATCTCAGCTGCACAACTCTACTAGCCGCCAAATTCAGCAGGAAGCAGTTAGAGC	7200	
Qy	7201	GGTSGTCGGCCAACTCCCAACAGCACTTAGGTTTTCTGTTGAGATGGGGACTCAGA	7260	
Db	7201	GGTSGTCGGCCAACTCCCAACAGCACTTAGGTTTTCTGTTGAGATGGGGACTCAGA	7260	
Qy	7261	GACAGGACTAGCTGGATTTCTTAGGCTGAYTAAGAAATCCYTAAGCCTAGSTGGGAAGGTG	7320	
Db	7261	GACAGGACTAGCTGGATTTCTTAGGCTGAYTAAGAAATCCYTAAGCCTAGSTGGGAAGGTG	7320	
Qy	7321	ACCAATCCACTTTAAACACGGGGCTTGCACTTAGYTACACTGACCAATCAGAGAG	7380	
Db	7321	ACCAATCCACTTTAAACACGGGGCTTGCACTTAGYTACACTGACCAATCAGAGAG	7380	
Qy	7381	CTCACTAAAAATGCTAATTTAGCAAGACAGGAGTAAAGAAATAGCCAAATCATYTTATTCG	7440	
Db	7381	CTCACTAAAAATGCTAATTTAGCAAGACAGGAGTAAAGAAATAGCCAAATCATYTTATTCG	7440	
Qy	7441	MTGAGCAGCAGCAGGGGCAATGATCGGGATATAAACCCAAAGTYTTTCGAGCGGCA	7500	
Db	7441	MTGAGCAGCAGCAGGGGCAATGATCGGGATATAAACCCAAAGTYTTTCGAGCGGCA	7500	
Qy	7501	CGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCAC	7560	
Db	7501	CGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCAC	7560	
Qy	7561	TCTATTAAATCTTCARCTGCR	7582	
Db	7561	TCTATTAAATCTTCARCTGCR	7582	
RESULT	5			
LOCUS	BD196244			
DEFINITION	Endogenous retroviral sequences, associated with autoimmune diseases and/or with pregnancy disorders.			
ACCESSION	BD196244			

RESULT 5  
BD196244  
LOCUS  
DEFINITI  
ACCESSION



```

VERSION      BD196244.1 GI:33006014
KEYWORDS     JP 2002512530-A/10.
SOURCE       unidentified
ORGANISM     unclassified sequences.
REFERENCE    1 (bases 1 to 2372)
AUTHORS      Beseme,F., Blond,J.L., Bouton,O., Mandrand,B. and Mallet,F.
TITLE        Endogenous retroviral sequences, associated with autoimmune
JOURNAL      diseases and/or with pregnancy disorders
              Patent: JP 2002512530-A 10 23-APR-2002;
COMMENT      BIO MERIEUX
              OS Unidentified
              PN JP 2002512530-A/10
              PD 23-APR-2002
              PF 06-JUL-1998 JP 1999508244
              PR 07-JUL-1997 FR 97/08815
              PI FREDERIC BESEME,JEAN LUC BLOND,OLIVIER BOUTON,BERNARD
              MANDRAND,
              PI FRANCOIS MALLET
              PC C12N15/48,C07K14/15,C12Q1/68,C07K16/10,C01N33/569 CC
              Strandedness: Single;
              CC Topology: Linear;
              CC Endogenous retroviral sequences, associated with autoimmune
              CC and/or with pregnancy disorders
              CC diseases
              FH Key Location/Qualifiers
              FT source 1..2372
              FT Location/Qualifiers
              1..2372
              /organism="unidentified"
              /mol_type="genomic DNA"
              /db_xref="taxon:32644"

ORIGIN
Query Match 11.7%; Score 888; DB 2; Length 2372;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 988; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1507 GGAAGGGAGAAATCTACTGCTTCTTGGAGAGACTAAGGAGGCATTGAGGAAGCG 1566
DB 1383 GGAAGGGAGAAATCTACTGCTTCTTGGAGAGACTAAGGAGGCATTGAGGAAGCG 1442
QY 1567 TGCCTCTGTACCTGACTCTTCTGAAGCCCACTAATCTTAAAGCGTAAGTTATCAC 1626
DB 1443 TGCCTCTGTACCTGACTCTTCTGAAGCCCACTAATCTTAAAGCGTAAGTTATCAC 1502
QY 1627 TCAGTCAGCTGCAGACATTAGAAAACCTTCAAAAGTCTGCCGTAGGCCCGGAGCAAC 1686
DB 1503 TCAGTCAGCTGCAGACATTAGAAAACCTTCAAAAGTCTGCCGTAGGCCCGGAGCAAC 1562
QY 1687 TTAGAAACCTATTGAACTTGGCACTCGGTTTATTAATAGAGATCAGGAGGAGCAG 1746
DB 1563 TTAGAAACCTATTGAACTTGGCACTCGGTTTATTAATAGAGATCAGGAGGAGCAG 1622
QY 1747 GCGGAACAGACAAACGGGATTAATAAGGCCACCGCTTTAGTCATGACCTCAGGCA 1806
DB 1623 GCGGAACAGACAAACGGGATTAATAAGGCCACCGCTTTAGTCATGACCTCAGGCA 1682
QY 1807 AGTGGACTTTGGAGGCTCTGGAAGGGAAGAGCTGGGCAATTAATGATCCCTAATAGGC 1866
DB 1683 AGTGGACTTTGGAGGCTCTGGAAGGGAAGAGCTGGGCAATTAATGATCCCTAATAGGC 1742
QY 1867 TTGCTTCCAGTCGGGCTTCAAGGACACTTTAAAAAGGATTTGCAAGTAGAAGTAAGCC 1926
DB 1743 TTGCTTCCAGTCGGGCTTCAAGGACACTTTAAAAAGGATTTGCAAGTAGAAGTAAGCC 1802
QY 1927 GCCCTTCGTCCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCGCCAGGGG 1986
DB 1803 GCCCTTCGTCCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCGCCAGGGG 1862
QY 1987 ACAAGGCTTTTGTAGTCAGAGCCCACTAACAGATGATCCAGCAGCAGACTGAGGGTG 2046

```

```

Db 1863 ACAAGGCTTTTGTAGTCAGAGCCCACTAACAGATGATCCAGCAGCAGACTGAGGGTG 1922
QY 2047 CTTGGGGCAAGGCCCATCCCATGCCATCACCTTCACAGAGCCCTGGGTATGCTTGACCAT 2106
Db 1923 CTTGGGGCAAGGCCCATCCCATGCCATCACCTTCACAGAGCCCTGGGTATGCTTGACCAT 1982
QY 2107 TGAGGGCCAGGAAGGTGTTCTCTCTGACACTGCTGGTGGGTCTTCTTAGTCTTACTTCTG 2166
Db 1983 TGAGGGCCAGGAAGGTGTTCTCTCTGACACTGCTGGTGGGTCTTCTTAGTCTTACTTCTG 2042
QY 2167 TCCCGGACAACTGTCTCTCAGATCTCTCACTATTCTGAGGGGTCCNTAAGAGCGGCAGT 2226
Db 2043 TCCCGGACAACTGTCTCTCAGATCTCTCACTATTCTGAGGGGTCCNTAAGAGCGGCAGT 2102
QY 2227 CACTAGATCTTTTCTCCAGCCCACTAAGTTATGAACCTGGGGAGCTTTATTTCTTTTCCACAT 2286
Db 2103 CACTAGATCTTTTCTCCAGCCCACTAAGTTATGAACCTGGGGAGCTTTATTTCTTTTCCACAT 2162
QY 2287 GCTTTTCTAATTATGCTTGAAGCCCACTACCTTCTTGAAGGAGACATTTCTAGCAAAA 2346
Db 2163 GCTTTTCTAATTATGCTTGAAGCCCACTACCTTCTTGAAGGAGACATTTCTAGCAAAA 2222
QY 2347 GCAGGGGCCATTATACCTGAACATAGGAGAGGAACACCCGTTGTGTNCCCTGCT 2406
Db 2223 GCAGGGGCCATTATACCTGAACATAGGAGAGGAACACCCGTTGTGTNCCCTGCT 2282
QY 2407 TGAGGAAGGAATTAATCTGAACTCTGGCAACAGAAAGACAAATATGAGCAGCCAAAGA 2466
Db 2283 TGAGGAAGGAATTAATCTGAACTCTGGCAACAGAAAGACAAATATGAGCAGCCAAAGA 2342
QY 2467 ATGCCCGTCTCTTCAAGTTAAACTAAAGG 2496
Db 2343 ATGCCCGTCTCTTCAAGTTAAACTAAAGG 2372

RESULT 6
BD267486
LOCUS
DEFINITION Endogenous nucleic fragment associated with an autoimmune disease,
              marking method and reagent.
ACCESSION BD267486
VERSION BD267486.1 GI:33077254
KEYWORDS JP 2002534980-A/27.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Homnidae; Homo.
REFERENCE 1 (bases 1 to 2372)
AUTHORS Baccala,G.P., Mallet,F. and Voisset,C.
TITLE Endogenous nucleic fragment associated with an autoimmune disease,
              marking method and reagent
JOURNAL Patent: JP 2002534980-A 27 22-OCT-2002;
COMMENT BIO MERIEUX
              OS Homo sapiens (human)
              PN JP 2002534980-A/27
              PD 22-OCT-2002
              PF 21-JAN-2000 JP 2000594929
              PR 21-JAN-1999 FR 99/00888
              PI GLAUCIA PARANHOS BACCALA,FRANCOIS MALLET,CECILE VOISSET PC
              C12N15/09,C07K14/15,C12Q1/02,C12Q1/68,G01N33/53,G01N33/53, PC
              G01N33/564,
              PC G01N33/566,G01N33/569,G01N33/58//C12P21/02,C12N15/00 CC
              Endogenous nucleic fragment associated with an autoimmune CC
              disease,
              CC marking method and reagent
              FH Key Location/Qualifiers
              FT source 1..2372
              FT /organism='Homo sapiens (human)'
              FT Location/Qualifiers
              1..2372
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              FEATURES
              source

```

[illegible]

Qy	2167	TCCCGGACAACTGTCTCCAGATCTGTCACTATTTCTGAGGGGTCCNTAAGACGGGCAGT	2226
Db	2043	TCCCGGACAACTGTCTCCAGATCTGTCACTATTTCTGAGGGGTCCNTAAGACGGGCAGT	2102
Qy	2227	CAC TAGATAC TTTTTC CAGCCACTAAAGTTATGAATCTGGGAGACTTTATTTCTTTTTCACAT	2286
Db	2103	CAC TAGATAC TTTTTC CAGCCACTAAAGTTATGAATCTGGGAGACTTTATTTCTTTTTCACAT	2162
Qy	2287	GC TTTTCTAATATATGCTTGAAAGCCCACTACCTTGTTTAGGGAGAGACATTTCTAGCAAAA	2346
Db	2163	GC TTTTCTAATATATGCTTGAAAGCCCACTACCTTGTTTAGGGAGAGACATTTCTAGCAAAA	2222
Qy	2347	GCAGGGCCATTATACACCTTGAAACATAGGAGAGAACACCCGTTTGCTGTCGTCCTGCT	2406
Db	2223	GCAGGGCCATTATACACCTTGAAACATAGGAGAGAGAACACCCGTTTGCTGTCGTCCTGCT	2282
Qy	2407	TGAGGAAGGAATTAATCCTGAAGTCTGGGCACACAGAAAGGACAATATATGGAGCGCAAAAGA	2466
Db	2283	TGAGGAAGGAATTAATCCTGAAGTCTGGGCACACAGAAAGGACAATATATGGAGCGCAAAAGA	2342
Qy	2467	ATGCCGCTCCTGTTCAAGTTAAACTAAAGG	2496
Db	2343	ATGCCGCTCCTGTTCAAGTTAAACTAAAGG	2372
RESULT 8			
LOCUS	AX027479	2372 bp	DNA linear PAT 16-SEP-2000
DEFINITION	Sequence 29 from Patent FR2788784.		
ACCESSION	AX027479		
VERSION	AX027479.1 GI:10188443		
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1		
AUTHORS	Mallet, F., Voisset, C. and Paranhos, B. G.		
JOURNAL	Patent: FR 2788784-A 29 28-JUL-2000;		
FEATURES	BIO MERIEUX (FR)		
source	Location/Qualifiers		
ORIGIN	1..2372		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
/db_xref="taxon:9606"			
Query Match 11.7%; Score 888; DB 2; Length 2372;			
Best Local Similarity 99.8%; Pred. No. 0;			
Matches 988; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	1507	GGGAAGGGAAGAAAACTCTACTGCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCG	1566
Db	1383	GGGAAGGGAAGAAAACTCTACTGCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGCG	1442
Qy	1567	TGCTCTCTCTGACCTGACTCTTCTGAGGCCAACTAATCTTAAAGCTTAAGTTTATCAC	1626
Db	1443	TGCTCTCTCTGACCTGACTCTTCTGAGGCCAACTAATCTTAAAGCTTAAGTTTATCAC	1502
Qy	1627	TCAGTCAGCTGCAGACATTAGAAAAAACTTTCAAAGTCTGCCGTAGGCCCGGAGCAAAAC	1686
Db	1503	TCAGTCAGCTGCAGACATTAGAAAAAACTTTCAAAGTCTGCCGTAGGCCCGGAGCAAAAC	1562
Qy	1687	TTAGAAAACCTATTTGAACCTTGGCAACTCGGTTTTTTTATATAGATCAGGAGGAGCAG	1746
Db	1563	TTAGAAAACCTATTTGAACCTTGGCAACCTCGTTTTTTTATATAGATCAGGAGGAGCAG	1622
Qy	1747	GC GGACAGACAAAACGGGATTAAAAAAGGCCACCGCTTTAGTATGATCCCTCAGGCA	1806
Db	1623	GC GGACAGACAAAACGGGATTAAAAAAGGCCACCGCTTTAGTATGATCCCTCAGGCA	1682

Qy	1807	AGTGGACTTTGGAGGCTCTGAAAGAGGAAAGCTGGGCAAAATTGAATGCTCTAATAGGGC	1861
Db	1683	AGTGGACTTTGGAGGCTCTGAAAGAGGAAAGCTGGGCAAAATTGAATGCTCTAATAGGGC	1742
Qy	1867	TTGCTTCCAGTGGCGTCTACAAGGACACTTTTAAAAAAGATTGTTCCAAGTATAGAGTAAGCC	1926
Db	1743	TTGCTTCCAGTGGCGTCTACAAGGACACTTTTAAAAAAGATTGTTCCAAGTATAGAGTAAGCC	1802
Qy	1927	GCCCTTTGCTCCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGG	1986
Db	1803	GCCCTTTGCTCCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGG	1862
Qy	1987	ACAAAGTCTTTTGAGTCAGAGCCACTAACCCAGATGATCCAGAGCAGAGCTGAGGGTG	2046
Db	1863	ACAAAGTCTTTTGAGTCAGAGCCACTAACCCAGATGATCCAGAGCAGAGCTGAGGGTG	1922
Qy	2047	CCTGGGCAAGCGCATCCCATGCGCATCACCCCTCACAGAGCCCTGGGTATGCTTGACCAT	2106
Db	1923	CCTGGGCAAGCGCATCCCATGCGCATCACCCCTCACAGAGCCCTGGGTATGCTTGACCAT	1982
Qy	2107	TGAGGGCCAGGAAGTTGTCTCTGGGACACTGTGTGGGTCTTCTTATGCTTACTCTTCTG	2166
Db	1983	TGAGGGCCAGGAAGTTGTCTCTGGGACACTGTGTGGGTCTTCTTATGCTTACTCTTCTG	2042
Qy	2167	TCCCGGCAACTGTCTCCAGATCTGTCACTATTCAGAGGGGTGCNTAAGACGGGCAGT	2226
Db	2043	TCCCGGCAACTGTCTCCAGATCTGTCACTATTCAGAGGGGTGCNTAAGACGGGCAGT	2102
Qy	2227	CACCTAGATCTTTTCCACGCCACTAGTTATGAACCTGGGGAGCTTTATTCCTTTTTCACAT	2286
Db	2103	CACCTAGATCTTTTCCACGCCACTAGTTATGAACCTGGGGAGCTTTATTCCTTTTTCACAT	2162
Qy	2287	GCTTTTCTAATTATGCTTGAAGCCCCACCTACCTTGTTAGGGAGAGACATTTCTAGCAAAA	2346
Db	2163	GCTTTTCTAATTATGCTTGAAGCCCCACCTACCTTGTTAGGGAGAGACATTTCTAGCAAAA	2222
Qy	2347	GCAGGGGCCAATTATACACCTGAACATAGGAGAGGAAACACCCGTTTGTGTGTCCTGCT	2406
Db	2223	GCAGGGGCCAATTATACACCTGAACATAGGAGAGGAAACACCCGTTTGTGTGTCCTGCT	2282
Qy	2407	TGAGGAGGAATTAATCTCGAAGCTGGGCAACAGAGGACATATGAGAGGCCAAAGA	2466
Db	2283	TGAGGAGGAATTAATCTCGAAGCTGGGCAACAGAGGACATATGAGAGGCCAAAGA	2342
Qy	2467	ATGCCCTGCTGTTTCAAGTTTAAACTAAAGG	2496
Db	2343	ATGCCCTGCTGTTTCAAGTTTAAACTAAAGG	2372
RESULT 9			
AF072504		2372 bp	linear
LOCUS	AF072504	2372 bp	linear
DEFINITION	Homo sapiens endogenous retrovirus W sequence.		
ACCESSION	AF072504		
VERSION	AF072504.1		
KEYWORDS	GI:4262288		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1. (bases 1 to 2372) Blond, J.L., Beseme, F., Duret, L., Bouton, O., Bedin, F., Perron, H., Mandrand, B. and Mallet, F.		
TITLE	Molecular characterization and placental expression of HERV-W, a new human endogenous retrovirus family		
JOURNAL	J. Virol. 73 (2), 1175-1185 (1999)		
PUBMED	9882319		
REFERENCE	2. (bases 1 to 2372) Blond, J.-L., Beseme, F. and Mallet, F.		
AUTHORS	Direct Submission		
TITLE	Submitted (19-JUN-1998) UM103 CNRS-biomevieux, ENS Lyon, 46 allée d'Italie, Lyon, Cedex 07 69364, France		
JOURNAL			

```

FEATURES
    source          Location/Qualifiers
                    1..2372
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="cl.44.4"
                    /tissue_type="placenta"
                    /clone_lib="Clontech 5'-stretch plus library, Cat number
                    HU5014a"
                    1..2372
                    /note="corresponds to gag and pol"
                    repeat_region 1..2372
                    /rpt_family="HERV-W"
                    5'UTR <1..6
                    /note="R"
                    5'UTR 7..460
                    /note="US"
                    primer_bind 464..481
                    /note="putative primer binding site for tRNA-W"

ORIGIN
Query Match 11.7%; Score 888; DB 5; Length 2372;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 988; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1507 GGGAGGGAAGAAATCCTACTGCTTCTGGAGACATAAGGAGGCATTGAGGAGCG 1566
DB 1383 GGGAGGGAAGAAATCCTACTGCTTCTGGAGACATAAGGAGGCATTGAGGAGCG 1442
QY 1567 TGCCTCTCTGTCACCTGACTCTTCTGAAGCCAACTTAATCTTAAAGCGTAAGTTTATCAC 1626
DB 1443 TGCCTCTCTGTCACCTGACTCTTCTGAAGCCAACTTAATCTTAAAGCGTAAGTTTATCAC 1502
QY 1627 TCAGTCAGCTGCAGACATATAGAAAAAATCTCAAAGTCTGCCGTAGGCCCGGAGCAAAAC 1686
DB 1503 TCAGTCAGCTGCAGACATATAGAAAAAATCTCAAAGTCTGCCGTAGGCCCGGAGCAAAAC 1562
QY 1687 TTAGAAACCTATTGAACCTTGGAACCTGCGTTTATTAATAGATCAGAGGAGCGAG 1746
DB 1563 TTAGAAACCTATTGAACCTTGGAACCTGCGTTTATTAATAGATCAGAGGAGCGAG 1622
QY 1747 GCGGAACAGACAAAACGGGATTAAAAAAGGCCACCGCTTTAGTCATGACCTCAGGCA 1806
DB 1623 GCGGAACAGACAAAACGGGATTAAAAAAGGCCACCGCTTTAGTCATGACCTCAGGCA 1682
QY 1807 AGTGGACTTTGGAGGCTCTGGAAAAGGAAAAGCTGGGCAAAATTGAATCCCTAATAGGCG 1866
DB 1683 AGTGGACTTTGGAGGCTCTGGAAAAGGAAAAGCTGGGCAAAATTGAATCCCTAATAGGCG 1742
QY 1867 TTGCTTCCAGTGGCGGTCTCAGGACACCTTTAAAAAAGATTGTCCAGTAGAAGTAAGCC 1926
DB 1743 TTGCTTCCAGTGGCGGTCTCAGGACACCTTTAAAAAAGATTGTCCAGTAGAAGTAAGCC 1802
QY 1927 GCCCTCTGTCGATGCCCTTATTTCAAGGGAATCACTGGAAGGCCACTGCCCCAGGGG 1986
DB 1803 GCCCTCTGTCGATGCCCTTATTTCAAGGGAATCACTGGAAGGCCACTGCCCCAGGGG 1862
QY 1987 ACAAAAGTCTTTTGAAGTCAGAACCCACTAAACAGATGATCCAGCAGCAGGACTGAGGGTG 2046
DB 1863 ACAAAAGTCTTTTGAAGTCAGAACCCACTAAACAGATGATCCAGCAGCAGGACTGAGGGTG 1922
QY 2047 CTGGGGGCAAGCCCATCCCATGATCACTCCATCAGAGCCCTGGGTATGCTTGACCAT 2106
DB 1923 CTGGGGGCAAGCCCATCCCATGATCACTCCATCAGAGCCCTGGGTATGCTTGACCAT 1982
QY 2107 TGAGGGCCAGGAAGTTGTCTCTGGACACTGGTGGGTCTTTAGTCTTACTCTTCTG 2166
DB 1983 TGAGGGCCAGGAAGTTGTCTCTGGACACTGGTGGGTCTTTAGTCTTACTCTTCTG 2042
QY 2167 TCCCGGACAACTGCTCCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGACGGGCAGT 2226
DB 2043 TCCCGGACAACTGCTCCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGACGGGCAGT 2102

```

```

QY 2227 CACTAGACTATTTTCCAGCCACTTAAGTTATGAATCGGAGCTTTATTTTTCACAT 2286
DB 2103 CACTAGACTATTTTCCAGCCACTTAAGTTATGAATCGGAGCTTTATTTTTCACAT 2162
QY 2287 GCTTTTCTAATATGCTTTGAAAGCCCACTACCTTTGTTAGGAGAGACATTTCTAGCAAAA 2346
DB 2163 GCTTTTCTAATATGCTTTGAAAGCCCACTACCTTTGTTAGGAGAGACATTTCTAGCAAAA 2222
QY 2347 GCAGGGGCATATATACCTGAACATAGGAGAAAGAACACCCGTTGTTGTTNCCCTGCT 2406
DB 2223 GCAGGGGCATATATACCTGAACATAGGAGAAAGAACACCCGTTGTTGTTNCCCTGCT 2282
QY 2407 TGAGGAAGGAATTAATCTGAACTCTGGCAACAGAGGACAAATATGACGAGCCAAAGA 2466
DB 2283 TGAGGAAGGAATTAATCTGAACTCTGGCAACAGAGGACAAATATGACGAGCCAAAGA 2342
QY 2467 ATGCCCGCTCTGTTCAAGTTAAACTAAAGG 2496
DB 2343 ATGCCCGCTCTGTTCAAGTTAAACTAAAGG 2372

RESULT 10
BD221806
LOCUS          2599 bp      DNA      linear      PAT 17-JUL-2003
DEFINITION    Nucleic sequence and deduced protein sequence family with human
               endogenous retroviral motifs, and their uses.
ACCESSION     BD221806
VERSION       BD221806.1  GI:33031576
KEYWORDS      JP 2002518051-A/1.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
               Hominoidea; Homo.
REFERENCE     1 (bases 1 to 2599)
AUTHORS      Alliel,P.M., Perin,J.P. and Rieger,F.
TITLES       Nucleic sequence and deduced protein sequence family with human
               endogenous retroviral motifs, and their uses
JOURNAL       Patent: JP 2002518051-A 1 25-JUN-2002;
               INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE
COMMENT       OS Homo sapiens (human)
               PN JP 2002518051-A/1
               PD 25-JUN-2002
               PP 23-JUN-1999  JP 2000556036
               PR 23-JUN-1998  FR 98/07920
               PI PATRICK M ALLIEL, JEAN PIERRE PERIN, FRANCOIS RIEGER PC
               C12N15/09,A01K67/027,A61K31/711,A61K39/21,A61K48/00,A61P21/00, PC
               A61P25/00.
               PC A61P37/06,C07K14/15,C12Q1/68,C12Q1/70,C12N15/00 CC Nucleic
               sequence and deduced protein sequence family with CC
               human endogenous
               CC retroviral motifs, and their uses
               FH Key      Location/Qualifiers
               FT source  1..2599
               FT          /organism='Homo sapiens (human)'.
               FEATURES
               source
               1..2599
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"

ORIGIN
Query Match 10.3%; Score 779; DB 2; Length 2599;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 4802 AGGCCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAATAAGGCCTAGTTTCATGAAA 4861
DB 151 AGGCCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAATAAGGCCTAGTTTCATGAAA 210
QY 4862 TAAATCCCAAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTCCAGG 4921
DB 211 TAAATCCCAAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTCCAGG 270

```

QY 4922 CCACAGTAAACCCAGGAGTATCCAGGGTGTAGGTATACGATATCACTTACACTGGCCCT 4981  
DB |||||  
DB 271 CCACAGTAAACCCAGGAGTATCCAGGGTGTAGGTATACGATATCACTTACACTGGCCCT 330  
QY 4982 GAAGGCCACAGTCTCTCAGGGAAGGTTCGAGAAAAATGAATGAAVACTCAAAAGGACATCTAA 5041  
DB |||||  
DB 331 GAAGGCCACAGTCTCTCAGGGAAGGTTCGAGAAAAATGAATGAAVACTCAAAAGGACATCTAA 390  
QY 5042 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCCCTGTCTGTGGCCCTATAGCCCTTAAAAA 5101  
DB |||||  
DB 391 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCCCTGTCTGTGGCCCTATAGCCCTTAAAAA 450  
QY 5102 GAATCTGCAATTTTCCCAAAAAAGCAGGACTTAGCCCATACGAAATCTGTATGGAGGC 5161  
DB |||||  
DB 451 GAATCTGCAATTTTCCCAAAAAAGCAGGACTTAGCCCATACGAAATCTGTATGGAGGC 510  
QY 5162 CCTTCATPAACCAATGACCTTGTGCTTGACCCAGACAGACCACTTAGTTCGACAGCATCAC 5221  
DB |||||  
DB 511 CCTTCATPAACCAATGACCTTGTGCTTGACCCAGACAGACCACTTAGTTCGACAGCATCAC 570  
QY 5222 CTCCTTAGCAAAATATCAAGTTCTTAAACATTAAGGAACTTATCCCTGAGAGA 5281  
DB |||||  
DB 571 CTCCTTAGCAAAATATCAAGTTCTTAAACATTAAGGAACTTATCCCTGAGAGA 630  
QY 5282 GGGAAAGAACTATTCACCCCTGTGACATGGTATTAAGTCAAGTCCCTTCCCTCTAATTC 5341  
DB |||||  
DB 631 GGGAAAGAACTATTCACCCCTGTGACATGGTATTAAGTCAAGTCCCTTCCCTCTAATTC 690  
QY 5342 CCCATCCCTAGATACATCTCTGGGAAGGACCTACCCAGTCAATTTATATACCCCACTGC 5401  
DB |||||  
DB 691 CCCATCCCTAGATACATCTCTGGGAAGGACCTACCCAGTCAATTTATACCCCACTGC 750  
QY 5402 GGTAAAGTGGCTGGAGTCTTGGATACATCACTTGAGTCAAAATCTCTGGTACT 5461  
DB |||||  
DB 751 GGTAAAGTGGCTGGAGTCTTGGATACATCACTTGAGTCAAAATCTCTGGTACT 810  
QY 5462 GCCAAGGAACCTGAAATCCAGAGACAACGGTAGCTATCTGTGAACTCTTAGAGGA 5521  
DB |||||  
DB 811 GCCAAGGAACCTGAAATCCAGAGACAACGGTAGCTATCTGTGAACTCTTAGAGGA 870  
QY 5522 TTTCGGCTCTCTTCAAAACAACACAGGAGAAAGTAACTAAATCAATAATCCCCA 5581  
DB |||||  
DB 871 TTTCGGCTCTCTTCAAAACAACACAGGAGAAAGTAACTAAATCAATAATCCCCA 929  
QY 5582 TGSSCCTCCCTTATCATATTTTCTCTKTASTGTSTTTTACCTCTTTTCACTCTCACTG 5641  
DB |||||  
DB 930 TGSSCCTCCCTTATCATATTTTCTCTTACTGTCTTTTACCTCTTTCACTCTCACTG 989  
QY 5642 CACCCCTCCATGCGCTGTATGACCAAGTACGCTCCCTTACCMAGAGTTTCTATGAGAA 5701  
DB |||||  
DB 990 CACCCCTCCATGCGCTGTATGACCAAGTACGCTCCCTTACCAAGAGTTTCTATGAGAA 1049  
QY 5702 TGACAGGTCGCGGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCA 5761  
DB |||||  
DB 1050 TGACAGGTCGCGGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCA 1109  
QY 5762 CTTTCACTGCCACACCCATATGATGCCCGCACTGCTATCACTCTGCCACTCTTTGATGC 5821  
DB |||||  
DB 1110 CTTTCACTGCCACACCCATATGATGCCCGCACTGCTATCACTCTGCCACTCTTTGATGC 1169  
QY 5822 ATGCAAACTACTTATTTGACAGGAAAAATGATTATCTAGTTGCTCTGGAGGACTTG 5881  
DB |||||  
DB 1170 ATGCAAACTACTTATTTGACAGGAAAAATGATTATCTAGTTGCTCTGGAGGACTTG 1229  
QY 5882 GAGTCACTGTCTGTGAGTACTTCAACCAAACTGATGTCTGATGGGGGTGAGTTC 5941  
DB |||||  
DB 1230 GAGTCACTGTCTGTGAGTACTTCAACCAAACTGATGTCTGATGGGGGTGAGTTC 1289  
QY 5942 AAGATCAGGCAAGAGAAAAATGTAAGAACTATCTCCCACTCACCGGGTACATG 6001  
DB |||||  
DB 1290 AAGATCAGGCAAGAGAAAAATGTAAGAACTATCTCCCACTCACCGGGTACATG 1349

QY 6002 GCACCTTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCTCGTACCC 6061  
DB |||||  
DB 1350 GCACCTTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCTCGTACCC 1409  
QY 6062 ATACTCGCCTGGTAAGCCTATTTAATACCAACCTCACTGGGCTCCATGAGGTCTCGGCC 6121  
DB |||||  
DB 1410 ATACTCGCCTGGTAAGCCTATTTAATACCAACCTCACTGGGCTCCATGAGGTCTCGGCC 1469  
QY 6122 AAAACCTTACTAATCTGTGTGATATGCTTCCCTCCCTGAACTTTCARGCCATATGTTTCAATCC 6181  
DB |||||  
DB 1470 AAAACCTTACTAATCTGTGTGATATGCTTCCCTCCCTGAACTTTCARGCCATATGTTTCAATCC 1529  
QY 6182 CTGTACTCTGAACCAATGAAACCACTTCAGCACAGAAATAACACACCTTCCGTTTGTAGTAG 6241  
DB |||||  
DB 1530 CTGTACTCTGAACCAATGAAACCACTTCAGCACAGAAATAACACACCTTCCGTTTGTAGTAG 1589  
QY 6242 GACCTCTTGTTCCTCAATSTGGAAATAACCATACCTCAAACTCCACCTGTGTAAAAATTTA 6301  
DB |||||  
DB 1590 GACCTCTTGTTCCTCAATCTGGAATAACCATACCTCAAACTCCACCTGTGTAAAAATTTA 1649  
QY 6302 GCAATACTACATACACAACTCCCAATGCAATGAGTGGGTAACTTCCTCCACACAAA 6361  
DB |||||  
DB 1650 GCAATACTACATACACAACTCCCAATGCAATGAGTGGGTAACTTCCTCCACACAAA 1709  
QY 6362 TAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTGCTCAGGCTATCGTTGTTTGA 6421  
DB |||||  
DB 1710 TAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTGCTCAGGCTATCGTTGTTTGA 1769  
QY 6422 ATGGCTCTTACAGAACTATGTGCTTCTCTCATTTCTAGTGGCCCTCATGTCATCTACA 6481  
DB |||||  
DB 1770 ATGGCTCTTACAGAACTATGTGCTTCTCTCATTTCTAGTGGCCCTCATGTCATCTACA 1829  
QY 6482 CTGAACAAGATTATACAGTTATGTATATCTAAGCCCGCAACAAAAGAGTACCCATTC 6541  
DB |||||  
DB 1830 CTGAACAAGATTATACAGTTATGTATATCTAAGCCCGCAACAAAAGAGTACCCATTC 1889  
QY 6542 TTCCCTTTTGTATAGAGCAGGAGTCTAGGTGCACTAGGTATCGGCATTTGGCGGTATCA 6601  
DB |||||  
DB 1890 TTCCCTTTTGTATAGAGCAGGAGTCTAGGTGCACTAGGTATCGGCATTTGGCGGTATCA 1949  
QY 6602 CACCTCTACTAGTTCTACTACAACTATCTCAAGAACTAATATGGGACATGGAACGGG 6661  
DB |||||  
DB 1950 CACCTCTACTAGTTCTACTACAACTATCTCAAGAACTAATATGGGACATGGAACGGG 2009  
QY 6662 TCGCCGACTCCCTGGTCACTTGCAGAGTCACTTAACTCCCTAGCAGCAGTATGCTTC 6721  
DB |||||  
DB 2010 TCGCCGACTCCCTGGTCACTTGCAGAGTCACTTAACTCCCTAGCAGCAGTATGCTTC 2069

RESULT 11  
AR699755

LOCUS AR699755 2599 bp DNA linear PAT 14-SEP-2005  
DEFINITION Sequence 1 from patent US 6919438.  
ACCESSION AR699755  
VERSION AR699755.1 GI:75205544  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 2599)  
AUTHORS Alliel,P.M., Perin,J.-P. and Rieger,F.  
TITLE Nucleic sequence and deduced protein sequence family with human  
endogenous retroviral motifs, and their uses  
JOURNAL Patent: US 6919438-A 1 19-JUL-2005;  
Institut National de la Sante et de la Recherche Medical-Inserm;  
Paris;  
FRX;

## FEATURES

Location/Qualifiers  
1..2599  
/organism="unknown"  
/mol\_type="genomic DNA"  
source  
ORIGIN

Query Match		10.3%; Score 779; DB 2; Length 2599;
Best Local Similarity		98.9%; Pred. No. 0;
Matches 1899; Conservative		0; Mismatches 20; Indels 1; Gaps 1;
QY	4802	AGGCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGACCTAGTTCATGAAA 4861
DB	151	AGGCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGACCTAGTTCATGAAA 210
QY	4862	TAAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTCCAGG 4921
DB	211	TAAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTCCAGG 270
QY	4922	CCACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTTACACTCGGCT 4981
DB	271	CCACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTTACACTCGGCT 330
QY	4982	GAAGGCCACAGTCTCCAGGGAAGGTCGAGAAAATGAATGAAAATCTCAAGAGGACATCTAA 5041
DB	331	GAAGGCCACAGTCTCCAGGGAAGGTCGAGAAAATGAATGAAAATCTCAAGAGGACATCTAA 390
QY	5042	AAAAGCAAAACCCAGGAAACCCACTCACATGGCGTGTCTGTGCTATAGGCTTAAAAA 5101
DB	391	AAAAGCAAAACCCAGGAAACCCACTCACATGGCGTGTCTGTGCTATAGGCTTAAAAA 450
QY	5102	GAATCTGCAACTTTCCTCCAAABAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGC 5161
DB	451	GAATCTGCAACTTTCCTCCAAABAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGC 510
QY	5162	CTTTCATAACCAATGACCTTGTGCTTACCCAGAGCAGCAACTTAGTTGACAGACATCAC 5221
DB	511	CTTTCATAACCAATGACCTTGTGCTTACCCAGAGCAGCAACTTAGTTGACAGACATCAC 570
QY	5222	CTCCTTAGCAATATCAACAAGTCTTAAACAATPACAAGGAACCTATCCCTGAGAGA 5281
DB	571	CTCCTTAGCAATATCAACAAGTCTTAAACAATPACAAGGAACCTATCCCTGAGAGA 630
QY	5282	GGGAAAGAACTATTCCACCWGTGACATGTTAGTCAAGTCCCTTCYCTTAATTC 5341
DB	631	GGGAAAGAACTATTCCACCWGTGACATGTTAGTCAAGTCCCTTCYCTTAATTC 690
QY	5342	CCCATCCCTAGATACATCTCGGGAAGGACCTACCAGTCACTTTATYATACCCCACTGC 5401
DB	691	CCCATCCCTAGATACATCTCGGGAAGGACCTACCAGTCACTTTATYATACCCCACTGC 750
QY	5402	GGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATACACTTGAGTCAAAATCTCGATCT 5461
DB	751	GGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATACACTTGAGTCAAAATCTCGATCT 810
QY	5462	GCCAAAGGAACCTGAAAATCCAGGAGACACCGCTAGCTATTCTGTGAACCTTAGAGA 5521
DB	811	GCCAAAGGAACCTGAAAATCCAGGAGACACCGCTAGCTATTCTGTGAACCTTAGAGA 870
QY	5522	TTTGCGCTGCTCTTCAAAACAACACAGGAGAAAGTAACTAAATAATCAATAATCCCCA 5581
DB	871	TTTGCGCTGCTCTTCAAAACAACACAGGAGAAAGTAACTAAATAATCAATAATCCCCA 929
QY	5582	TGSGCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCTSTTTTCACTCTCACTG 5641
DB	930	TGSGCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCTSTTTTCACTCTCACTG 989
QY	5642	CACCCCTCATGCGCTGTATGACCAAGTACTCCCTTACCMAGAGTTTCTATGAGAA 5701
DB	990	CACCCCTCATGCGCTGTATGACCAAGTACTCCCTTACCMAGAGTTTCTATGAGAA 1049
QY	5702	TGCGAGGTCGCGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCA 5761
DB	1050	TGCGAGGTCGCGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCA 1109
QY	5762	CTTCTACGCCACACCCATATGCCCGCAACTGCTATCACTGTCCACTCTTTGATGC 5821
DB	1110	CTTCTACGCCACACCCATATGCCCGCAACTGCTATCACTGTCCACTCTTTGATGC 1169
QY	5822	ATGCAAAATCTCATTTATGACAGGAAAATGATTAATCTAGTTGCTCTGGAGGACTTG 5881
DB	1170	ATGCAAAATCTCATTTATGACAGGAAAATGATTAATCTAGTTGCTCTGGAGGACTTG 1229
QY	5882	GAGTCACTGTCTGTGGACTTACCTTACCCAAAACCTGGTATGTCTGATGGGGTGGAGTTC 5941
DB	1230	GAGTCACTGTCTGTGGACTTACCTTACCCAAAACCTGGTATGTCTGATGGGGTGGAGTTC 1289
QY	5942	AAGATCAGGCAAGAGAAAACATGTAAAGAAAGTAACTCTCCAACTCACTACCCSGGGTACATG 6001
DB	1290	AAGATCAGGCAAGAGAAAACATGTAAAGAAAGTAACTCTCCAACTCACTACCCSGGGTACATG 1349
QY	6002	GCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAAATCTACATGAAAACCTCCGTACCC 6061
DB	1350	GCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAAATCTACATGAAAACCTCCGTACCC 1409
QY	6062	ATATCGGCTCGTAAAGCTATTTAATACCACTTCACTCGGGTCCATGAGGTCCTCGGCC 6121
DB	1410	ATATCGGCTCGTAAAGCTATTTAATACCACTTCACTCGGGTCCATGAGGTCCTCGGCC 1469
QY	6122	AAACCCCTACTAACTGTGTGGATATGCCCTCCCTGAACTTCARGGCATATGTTCAATCC 6181
DB	1470	AAACCCCTACTAACTGTGTGGATATGCCCTCCCTGAACTTCARGGCATATGTTCAATCC 1529
QY	6182	CTGTACCTGAAACAATGGAAACAATTTACGACAGAAAATAAACAACCACTTCCGTTTGTAGTAG 6241
DB	1530	CTGTACCTGAAACAATGGAAACAATTTACGACAGAAAATAAACAACCACTTCCGTTTGTAGTAG 1589
QY	6242	GACCTTGTGTTTCCAATSTGGAATAACCCATACCTCAAACTCACTCTGTGTAAATTTA 6301
DB	1590	GACCTTGTGTTTCCAATSTGGAATAACCCATACCTCAAACTCACTCTGTGTAAATTTA 1649
QY	6302	GCAATACCTACATACACAACCACTCCCAATGCATCAGTGGGTAACTCTCCCAACAAA 6361
DB	1650	GCAATACCTACATACACAACCACTCCCAATGCATCAGTGGGTAACTCTCCCAACAAA 1709
QY	6362	TAGTCTGCTTACCCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCGTTGTTGA 6421
DB	1710	TAGTCTGCTTACCCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCGTTGTTGA 1769
QY	6422	ATGGCTCTTCAAGATCTATGTGCTCTCATTTCTTAGTGCCCTCCCTATGRCCTATCA 6481
DB	1770	ATGGCTCTTCAAGATCTATGTGCTCTCATTTCTTAGTGCCCTCCCTATGRCCTATCA 1829
QY	6482	CTGAACAGATTTATACAGTATGTCTATCTTAAGCCCGCAACAAAAGAGTACCCATTC 6541
DB	1830	CTGAACAGATTTATACAGTATGTCTATCTTAAGCCCGCAACAAAAGAGTACCCATTC 1889
QY	6542	TTCCCTTTTATAGGACGAGTGTAGGTGCACTAGGTACTGGCATTTGGCGGTATCA 6601
DB	1890	TTCCCTTTTATAGGACGAGTGTAGGTGCACTAGGTACTGGCATTTGGCGGTATCA 1949
QY	6602	CAACCTCTACTCAGTCTTACTACAACTATCTCAAGAACTAAATGGGGAACATGGAAACGG 6661
DB	1950	CAACCTCTACTCAGTCTTACTACAACTATCTCAAGAACTAAATGGGGAACATGGAAACGG 2009
QY	6662	TGCGCGACTCCCTGGTCACTTGAAGATCACTTAATCTCCCTAGCAGCAGTGTCTTC 6721
DB	2010	TGCGCGACTCCCTGGTCACTTGAAGATCACTTAATCTCCCTAGCAGCAGTGTCTTC 2069
RESULT 12		
AX007978		
LOCUS		Sequence 1 from Patent WO9967395.
DEFINITION		AX007978
ACCESSION		AX007978.1 GI:9995675
VERSION		
KEYWORDS		Homo sapiens (human)
SOURCE		Homo sapiens
ORGANISM		Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
		Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
		Hominidae; Homo.
REFERENCE		1



AUTHORS		Perin,J.P., Rieger,F. and Alliel,P.M.
TITLE		Nucleic sequence and deduced protein sequence family with human
JOURNAL		endogenous retroviral motifs, and their uses
PATENT		Patent: WO 9967395-A 1 29-DEC-1999;
INST NAT		INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER
FRANCOIS		FRANCOIS Location/Qualifiers
FEATURES		1..2599
source		/organism="Homo sapiens"
ORIGIN		/mol_type="unassigned DNA"
		/db_xref="taxon:9606"
Query Match		10.3%; Score 779; DB 2; Length 2599;
Best Local Similarity		98.9%; Pred. No. 0;
Matches 1899;		Conservative 0; Mismatches 20; Indels 1; Gaps 1;
QY	4802	AGGCCCTTCCCTCTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGGCACTAGTTTCATGAAA 4861
DB	151	AGGCCCTTCCCTCTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGGCACTAGTTTCATGAAA 210
QY	4862	TAATTCGAGATTCCGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTCCAGG 4921
DB	211	TAATTCGAGATTCCGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTCCAGG 270
QY	4922	CCACAGTAAACCCAGGAGGATCCAGGCGTTAGGTATACGATATCACTTACACTGGCCCT 4981
DB	271	CCACAGTAAACCCAGGAGGATCCAGGCGTTAGGTATACGATATCACTTACACTGGCCCT 330
QY	4982	GAAGGCCACAGTCTCAGGGAAGGTCCAGAAAATGAATGAAAVACTCAAGAGGACATCTAA 5041
DB	331	GAAGGCCACAGTCTCAGGGAAGGTCCAGAAAATGAATGAAAVACTCAAGAGGACATCTAA 390
QY	5042	AAAAGCAAAACCCAGGAAACCCACTCACATGGCGCTGTCTGTTGCCCTATAGCCTTAAAAA 5101
DB	391	AAAAGCAAAACCCAGGAAACCCACTCACATGGCGCTGTCTGTTGCCCTATAGCCTTAAAAA 450
QY	5102	GAATCTGCAACTTTCCTCCCAAAAGAGGACTTAGCCCATAGCAAAATGCTGTATGGAAGGC 5161
DB	451	GAATCTGCAACTTTCCTCCCAAAAGAGGACTTAGCCCATAGCAAAATGCTGTATGGAAGGC 510
QY	5162	CCCTTCAACCAATGACCTTGTGCTTGACCCAGAGACAGCACTTAGTTGTCAGACATCAC 5221
DB	511	CCCTTCAACCAATGACCTTGTGCTTGACCCAGAGACAGCACTTAGTTGTCAGACATCAC 570
QY	5222	CTCCTTAGCCAAATATCAACAGTTCTTAAACATTTACAAGGAACCTATCCCTTGAGAAGA 5281
DB	571	CTCCTTAGCCAAATATCAACAGTTCTTAAACATTTACAAGGAACCTATCCCTTGAGAAGA 630
QY	5282	GGGAAAGAACTATTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTTCYCTCTAATTTC 5341
DB	631	GGGAAAGAACTATTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTTCYCTCTAATTTC 690
QY	5342	CCCATCCCTAGATACATCTGGGAAGGACCTTACCAGTCATTTTATATYACCCCACTGC 5401
DB	691	CCCATCCCTAGATACATCTGGGAAGGACCTTACCAGTCATTTTATATYACCCCACTGC 750
QY	5402	GGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAATCAAACTCTGGATACT 5461
DB	751	GGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAATCAAACTCTGGATACT 810
QY	5462	GCCAAAGGAACCTTGAAAATCCAGGAGCAACGCTAGCTATTCTCTGTGAAACCTCTAGAGGA 5521
DB	811	GCCAAAGGAACCTTGAAAATCCAGGAGCAACGCTAGCTATTCTCTGTGAAACCTCTAGAGGA 870
QY	5522	TTTGCCCTGCTCTTCAAAACAAACCCAGGAGAAAGTAACATAAATCAATAATCCCCCA 5581
DB	871	TTTGCCCTGCTCTTCAAAACAAACCCAGGAGAAAGTAACATAAATCAATAATCCCCCA 929
QY	5582	TGGSCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCTSTTTTCACTCTCACATG 5641
DB	930	TGGSCTCCCTTATCATATTTTCTCTTTACTGTTCTTTTACCCTCTTTTCACTCTCACATG 989

QY	5642	CACCCCTCCATGCGCTGTATGACAGTAGTCCCTTACCMAGAGTTTCTTATGAGAA 5701
DB	990	CACCCCTCCATGCGCTGTATGACAGTAGTCCCTTACCMAGAGTTTCTTATGAGAA 1049
QY	5702	TGCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCA 5761
DB	1050	TGCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCA 1109
QY	5762	CTTTCACTGCCACACCCCATATGCCCCCGCAACTGCTATCACTCTGCCACTCTTTGATGC 5821
DB	1110	CTTTCACTGCCACACCCCATATGCCCCCGCAACTGCTATCACTCTGCCACTCTTTGATGC 1169
QY	5822	ATGCAATATCTCATTTATGACAGGAAATATGTAATCTTAGTTGCTCGGAGGACTTG 5881
DB	1170	ATGCAATATCTCATTTATGACAGGAAATATGTAATCTTAGTTGCTCGGAGGACTTG 1229
QY	5882	GAGTCACTGCTGTGAGACTTACTTCAACCAAACTGATGCTGATGCGGGTGGAGTTTC 5941
DB	1230	GAGTCACTGCTGTGAGACTTACTTCAACCAAACTGATGCTGATGCGGGTGGAGTTTC 1289
QY	5942	AAGATCAGCAAGAGAAAAACATGTAAAGAAAGTAATCTCCCAACTCACCSGGGTACATG 6001
DB	1290	AAGATCAGCAAGAGAAAAACATGTAAAGAAAGTAATCTCCCAACTCACCSGGGTACATG 1349
QY	6002	GCACCTCTAGCCCTCAAGAGCTAGATCTCTCAAAAACATCATGAAACCTCCCGTACCC 6061
DB	1350	GCACCTCTAGCCCTCAAGAGCTAGATCTCTCAAAAACATCATGAAACCTCCCGTACCC 1409
QY	6062	ATACTCGCTGCTGAGCTTATTTAATACCACCTCACTGGGCTCCATGAGGTCTCGGCC 6121
DB	1410	ATACTCGCTGCTGAGCTTATTTAATACCACCTCACTGGGCTCCATGAGGTCTCGGCC 1469
QY	6122	AAAAACCTACTAATCTGTGGATATGCTCCCTGAACTTTCARGCCATATGTTTCAATCC 6181
DB	1470	AAAAACCTACTAATCTGTGGATATGCTCCCTGAACTTTCARGCCATATGTTTCAATCC 1529
QY	6182	CTGTACTGAAACATAGGAAACAACTTCAGCAAGAAATAAACACCACTCCGTTTTAGTAG 6241
DB	1530	CTGTACTGAAACATAGGAAACAACTTCAGCAAGAAATAAACACCACTCCGTTTTAGTAG 1589
QY	6242	GACCTCTGTTTTCCAAATSTGGAATAACCCATACCTCAAACTCCTGTAATAATTAA 6301
DB	1590	GACCTCTGTTTTCCAAATSTGGAATAACCCATACCTCAAACTCCTGTAATAATTAA 1649
QY	6302	GCAATACTACATACACCAACTCCCAATGATCAGGTGGGTAACTCTCCCAACAAA 6361
DB	1650	GCAATACTACATACACCAACTCCCAATGATCAGGTGGGTAACTCTCCCAACAAA 1709
QY	6362	TAGTCTGCCCTACCCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTGA 6421
DB	1710	TAGTCTGCCCTACCCCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTGA 1769
QY	6422	ATGGCTCTTCAGAACTATGCTGCTCTCTCAATCTTAGTGGCCCTCCTATGRCATCTACA 6481
DB	1770	ATGGCTCTTCAGAACTATGCTGCTCTCTCAATCTTAGTGGCCCTCCTATGRCATCTACA 1829
QY	6482	CTGAAACAGATTTATACAGTTATGTCATATCTAAGCCCGCAACAAAAGAGTACCCATTC 6541
DB	1830	CTGAAACAGATTTATACAGTTATGTCATATCTAAGCCCGCAACAAAAGAGTACCCATTC 1889
QY	6542	TTCTTTTGTATTAGGACGAGGTGCTAGGTGCACTAGGTACTGGCATTTGGCGGTATCA 6601
DB	1890	TTCTTTTGTATTAGGACGAGGTGCTAGGTGCACTAGGTACTGGCATTTGGCGGTATCA 1949
QY	6602	GAACTCTACTAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAAACGGG 6661
DB	1950	CAACCTCTACTAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAAACGGG 2009
QY	6662	TGCGCGACTCCCTGCTGCTTGGCAAGATCACTTAACTCCCTAGCAGCAGTAGTCTCTTC 6721
DB	2010	TGCGCGACTCCCTGCTGCTTGGCAAGATCACTTAACTCCCTAGCAGCAGTAGTCTCTTC 2069





```

QY 6602 CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGG 6661
Db 7870 CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGG 7929
QY 6662 TCGCCGACTCCCTGGTCACTTCCAGAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTTC 6721
Db 7930 TCGCCGACTCCCTGGTCACTTCCAGAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTTC 7989

RESULT 14
AY101582
LOCUS
DEFINITION
Homo sapiens isolate 21 endogenous retrovirus HERV-W, ERVW1 locus,
allele A, complete sequence.
ACCESSION
AY101582
VERSION
AY101582.1
KEYWORDS
GI:37544399
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE
1 (bases 1 to 10222)
Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.,
The endogenous retroviral locus ERVW1 is a bona fide gene involved
in hominoid placental physiology
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
14757826
2 (bases 1 to 10222)
Mallet,F., Bouton,O. and Oriol,G.
Direct Submission
Submitted (06-MAY-2002) Retrovirology Department, UMR 2142
CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
Location/Qualifiers
1..10222
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="21"
/isolation_source="PBMC of individual 21"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-q22"
/sex="female"
/cell_type="PBMC"
/enote="isolated by PCR; ERVW1 locus, allele A
endogenous_virus: HERV-W"
1..780
Join(248..811,2887..9813)
/note="putative mRNA transcript 1; gag-pol pseudogene"
Join(248..811,7585..9813)
/product="envelope glycoprotein"
/note="putative mRNA transcript 2"
Join(248..811,9248..9813)
/note="putative mRNA transcript 3; no coding region"
7812..9428
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAW68161.1"
/db_xref="GI:37544400"
/translation="MALPYHIEFLFTLLPSFTLTAPPCCRCMTSSSPYQEFLLRMWRP
GNIDAPYSRLSKGPTFTATHMPNRCVHSATLCMHANTHYMTCKMINPSCGGLGV
TVCWTYFTGTGSDGGVQDOQAREKHVEISQLTRVGTSSPYGLDLKLHETLRT
HRLVSLFNLTILGLHVSQAQNTICLPLNFRPVSIPVPEQNNFSTEINTSV
LVGLVSLNEIHTLSNLKVCFSNTTYTNSQIRWPTPTQIVCLPSGIFPVCGTSA
YRLNGSSEMSFLSFLVPMPTIYTEQDLYSVISKPRKRPILPFIAGVLGALG
TGIGTITSTQPYKLSQELNGDMERVADSLVTLQDLSLAALVQLNRRLDLLTAE
RGGTCLFLGECCQYVNSGIVTEKVKELRDKIQRARELRTGPGLLSOWMPLP
FLGPLAAIILLIFGPCIFNLLVNFVSSRIENAVKLQEPKMQSKTKIYRRPLDRPAS
RSDVNDYKGTPEEISAQPLLRPNAGSS"
9487..10222

```

LTR

## ORIGIN

```

Query Match 10.3%; Score 779; DB 5; Length 10222;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 4802 AGGCTTCCCTCTAGGACAGAAAAGGCCCAAGAGTAAATAAGGCACTAGTTTCATGAAA 4861
Db 7034 AGGCTTCCCTCTAGGACAGAAAAGGCCCAAGAGTAAATAAGGCACTAGTTTCATGAAA 7093
QY 4862 TAAATCCAGATTCGGGACTTCCCGAGGCTTACAGAGTGCATATAGCCCTGCTTTCCAGG 4921
Db 7094 TAAATCCAGATTCGGGACTTCCCGAGGCTTACAGAGTGCATATAGCCCTGCTTTCCAGG 7153
QY 4922 CCACAGTAAACCCAGGAGTATCCCGAGGCTTAGGTATACGATATCACTTACACTGGCCCT 4981
Db 7154 CCACAGTAAACCCAGGAGTATCCCGAGGCTTAGGTATACGATATCACTTACACTGGCCCT 7213
QY 4982 GAAGGCCACAGTCTCAGGGAAGGTGCAGAAAATGAATGAAYACTCAAAGGACATCTAA 5041
Db 7214 GAAGGCCACAGTCTCAGGGAAGGTGCAGAAAATGAATGAAYACTCAAAGGACATCTAA 7273
QY 5042 AAAAGCAAAACCCAGGAAACCCACCTCACATGCGCTGTCTGTTCCTATAGCCTTAATA 5101
Db 7274 AAAAGCAAAACCCAGGAAACCCACCTCACATGCGCTGTCTGTTCCTATAGCCTTAATA 7333
QY 5102 GAATCTGCACACTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 5161
Db 7334 GAATCTGCACACTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 7393
QY 5162 CTTTCAATAACCAATGACCTTGTGTGACCCAGAGCCAACTTAGTTAGTCAGACATCAC 5221
Db 7394 CTTTCAATAACCAATGACCTTGTGTGACCCAGAGCCAACTTAGTTAGTCAGACATCAC 7453
QY 5222 CTCCTTAGCCCAATATCAACAAGTCTTAAACATTTACAGGAACTATCCCTGAGAAGA 5281
Db 7454 CTCCTTAGCCCAATATCAACAAGTCTTAAACATTTACAGGAACTATCCCTGAGAAGA 7513
QY 5282 GGGAAAAGAACTATTCCACCCWMTGACATGTATTAGTCAAGTCCCTTCYCTCTAATTC 5341
Db 7514 GGGAAAAGAACTATTCCACCCWMTGACATGTATTAGTCAAGTCCCTTCYCTCTAATTC 7573
QY 5342 CCATCCCTTAGATATACATCTCTGGAAAGGCCCTTACCAGTCATTTATYATACCCCACTGC 5401
Db 7574 CCATCCCTTAGATATACATCTCTGGAAAGGCCCTTACCAGTCATTTATYATACCCCACTGC 7633
QY 5402 GGTAAAGTGGCTGGAGTGGAGTCTTGATATACATCACACTTACGACAAATCTCTGATCT 5461
Db 7634 GGTAAAGTGGCTGGAGTGGAGTCTTGATATACATCACACTTACGACAAATCTCTGATCT 7693
QY 5462 GCCAAAGAACTCGAAATCCAGGAGACAAACGCTAGCTATTCTGTGAACCTCTAGAGGA 5521
Db 7694 GCCAAAGAACTCGAAATCCAGGAGACAAACGCTAGCTATTCTGTGAACCTCTAGAGGA 7753
QY 5522 TTTGGCCTGCTCTTCAAAACAAACACAGGAGAAAGTAACTAAATCATATAATCCCCCA 5581
Db 7754 TTTGGCCTGCTCTTCAAAACAAACACAGGAGAAAGTAACTAAATCATATAATCCCCCA 7812
QY 5582 TGGSCCTCCCTTATCATATTTTCTTATSTGTTSTTTTACCTSTTTTCACTCTCACTG 5641
Db 7813 TGGSCCTCCCTTATCATATTTTCTTATSTGTTSTTTTACCTSTTTTCACTCTCACTG 7872
QY 5642 CACCCCTCCATCGCGCTGATGACAGTACCTCCCTVACOMAGAGTTCTATGAGAA 5701
Db 7873 CACCCCTCCATCGCGCTGATGACAGTACCTCCCTVACOMAGAGTTCTATGAGAA 7932
QY 5702 TGCAGCTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCA 5761
Db 7932 TGCAGCTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCA 7992
QY 5762 CCTTCACTGCCCCACACCCATATGCCCCGAACTGTCTATCACTCTGCCACTCTTTGATGC 5821
Db 7993 CCTTCACTGCCCCACACCCATATGCCCCGAACTGTCTATCACTCTGCCACTCTTTGATGC 8052

```

QY	5822	ATGCAATCTACTATTTGGACAGGAAAAATGATTAATCCTAGTTGTCTCTGGAGACTTG	5881
Db	8053	ATGCAATCTACTATTTGGACAGGAAAAATGATTAATCCTAGTTGTCTCTGGAGACTTG	8112
QY	5882	GAGTCACTGTCTGTTGGACTTACTTCCACCAACTGGTATGCTCATGGGGTGGAGTTC	5941
Db	8113	GAGTCACTGTCTGTTGGACTTACTTCCACCAACTGGTATGCTCATGGGGTGGAGTTC	8172
QY	5942	AAGATCAGGCAAGAGAAAAACATGTAAAGAAAGTAAATCTCCAACTCACCCSGGTACATG	6001
Db	8173	AAGATCAGGCAAGAGAAAAACATGTAAAGAAAGTAAATCTCCAACTCACCCSGGTACATG	8232
QY	6002	GCACCTCTAGCCCTTACAAAGACTAGATCTCTCAAAGACTACATGAACCCCTCGTACCC	6061
Db	8233	GCACCTCTAGCCCTTACAAAGACTAGATCTCTCAAAGACTACATGAACCCCTCGTACCC	8292
QY	6062	ATACTCGCTGCTGAAGCCTATTATTAACCAACCTCACTGGGCTCCATGAGGTCTCGGCC	6121
Db	8293	ATACTCGCTGCTGAAGCCTATTATTAACCAACCTCACTGGGCTCCATGAGGTCTCGGCC	8352
QY	6122	AAAAACCTACTAATCTGTTGGATATGCTTCCCTTGAACCTTCARGGCATATGTTCAATCC	6181
Db	8353	AAAAACCTACTAATCTGTTGGATATGCTTCCCTTGAACCTTCARGGCATATGTTCAATCC	8412
QY	6182	CTGTACTCTGAACAATGGAAACAACTTCAGACAGAAATAAACAACCACTCCGTTTAGTAG	6241
Db	8413	CTGTACTCTGAACAATGGAAACAACTTCAGACAGAAATAAACAACCACTCCGTTTAGTAG	8472
QY	6242	GACCTCTGTTTCCAAATGGAATAACCACTCAACCTCAACCTCAACCTGTTGTAATAATTA	6301
Db	8473	GACCTCTGTTTCCAAATGGAATAACCACTCAACCTCAACCTCAACCTGTTGTAATAATTA	8532
QY	6302	GCAATACTACATACACCAACTCCCAATGCATCAGGTGGGTAACTCTCTCCACACAAA	6361
Db	8533	GCAATACTACATACACCAACTCCCAATGCATCAGGTGGGTAACTCTCTCCACACAAA	8592
QY	6362	TAGTCTGCCTACCTCAGAAATATTTTGTCTGTGGTACCTCAGCTATCGTTGTTGA	6421
Db	8593	TAGTCTGCCTACCTCAGAAATATTTTGTCTGTGGTACCTCAGCTATCGTTGTTGA	8652
QY	6422	ATGGCTCTTCAGAACTATGCTCTCTCTCAATCTTAGTGCCCTCCATGRCATCTACA	6481
Db	8653	ATGGCTCTTCAGAACTATGCTCTCTCTCAATCTTAGTGCCCTCCATGRCATCTACA	8712
QY	6482	CTGAACAAGATTATACAGTTATGTCATATCTAAGCCCGCAACAAAAGAGTACCCATTC	6541
Db	8713	CTGAACAAGATTATACAGTTATGTCATATCTAAGCCCGCAACAAAAGAGTACCCATTC	8772
QY	6542	TTCCCTTTTGTATAGAGCAGAGTGTAGGTGCATAGGTACTGGCATTTGGCGGTATCA	6601
Db	8773	TTCCCTTTTGTATAGAGCAGAGTGTAGGTGCATAGGTACTGGCATTTGGCGGTATCA	8832
QY	6602	CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG	6661
Db	8833	CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG	8892
QY	6662	TGCGCGACTCCCTGGTTCACCTTGCAGATCAACTTAACTCCCTAGCAGAGTAGTCCCTTC	6721
Db	8893	TGCGCGACTCCCTGGTTCACCTTGCAGATCAACTTAACTCCCTAGCAGAGTAGTCCCTTC	8952
RESULT 15			
AY101585			
LOCUS			
DEFINITION			
Homo sapiens isolate 79 endogenous retrovirus HERV-W, ERVW1 locus, allele B, complete sequence.			
ACCESSION			
AY101585			
KEYWORDS			
AY101585.1 GI:37544405			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			

REFERENCE		Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS		1 (bases 1 to 10222)	
TITLE		Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.	
JOURNAL		The endogenous retroviral locus ERVW1 is a bona fide gene involved in hominoid placental physiology	
PUBMED		Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)	
REFERENCE		2 (bases 1 to 10222)	
AUTHORS		Mallet, F., Bouton, O. and Oriol, G.	
TITLE		Direct Submission	
JOURNAL		Submitted (06-MAY-2002) Retrovirology Department, UMR 2142 CNRS-biomerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France	
FEATURES		Location/Qualifiers	
source		1..10222	
		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/isolate="79"	
		/isolation_source="PBM of individual 79"	
		/db_xref="taxon:9606"	
		/chromosome="7"	
		/map="7q21-q22"	
		/sex="male"	
		/cell_type="PBM"	
		/note="caucasian"	
		isolated by PCR; ERVW1 locus, allele B	
		endogenous_virus: HERV-W"	
LTR		1..780	
mRNA		join(248..811,2887..9813)	
		/note="putative mRNA transcript 1; gag-pol pseudogene"	
mRNA		join(248..811,7585..9813)	
		/product="envelope glycoprotein"	
mRNA		/note="putative mRNA transcript 2"	
		join(248..811,9248..9813)	
		/note="putative mRNA transcript 3; no coding region"	
CDS		7812..9428	
		/note="syncytin"	
		/codon_start=1	
		/product="envelope glycoprotein"	
		/protein_id="AA068164.1"	
		/db_xref="GI:37544406"	
		/translation="MALPYHFLFTVLLPSFTLTAPPCCRCMTSSSPQEFRLMRQRP	
		GNIDAPSRSLKSGTPTTATHMPNCTHSATLCHMANTHWKMINPSCFGLGV	
		TVCYTFYTGMSDGGVQDQREKHVEVLSQLRVHGTSPPYKGLDLKLHETLRT	
		HTRVLSFNFTLTGLHVEYSAQNCTNCLPLNFRPVYSIPVEQNNFSTINTSV	
		LVGLVSNLEITHTSNLTCKVFSNTVYTNQSCIRWTPPTQIVCLPSGIFVCGTSA	
		YRLNGSSESCFLSFLVPMTIYEDOLYSVLSKPKRVPILPFVIGAGVLGALG	
		TGIGITTSFTFYKLSQELNGDMERVADSLVTLDQLNSLAADVVLQNRALDLTAE	
		RGGTCFLFGECCYYNQSGIVTEKVEIRDIRIQRRABELRNTGPGLLSQMPWILP	
		FLGLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRPLDRPASP	
		RSDVNDIKGTPPEISAAQPLLRPNAGSS"	
		9487..10222	
LTR		ORIGIN	
Query Match		10.3%; Score 779; DB 5; Length 10222;	
Best Local Similarity		98.9%; Pred. No. 0;	
Matches 1899;		Conservative 0; Mismatches 20; Indels 1; Gaps 1;	
QY	4802	AGGCTTCCCTCTAGGACAGAAAGCCCAAGAGGTAAAGGACCTAGTTCATGAAA	4861
Db	7034	AGGCTTCCCTCTAGGACAGAAAGCCCAAGAGGTAAAGGACCTAGTTCATGAAA	7093
QY	4862	TAAATCCAGATTCCGACTTCCCGAGGCTTACAGAGTCAACAATAGCCCTGCTTCCAGG	4921
Db	7094	TAAATCCAGATTCCGACTTCCCGAGGCTTACAGAGTCAACAATAGCCCTGCTTCCAGG	7153
QY	4922	CCACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATACGATATACATTACCTGGCCT	4981
Db	7154	CCACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATACGATATACATTACCTGGCCT	7213
QY	4982	GAAGCCACAGCTCCTCAGGGAAGTTCGAGAAAATCAATGAAAYACTCAAAGGACATCTAA	5041

Db 7214 GAAGGCCACAGTCTCTAGGGAGAGTTCGAGAAATGAATGAACACTCAAAGGACATCTAA 7273  
Qy 5042 AAAAGCAAAACCCAGGAAACCCACCTCACATGGGCTGTTCTGTTCCTATAGCCTTAAAAA 5101  
Db 7274 AAAAGCAAAACCCAGGAAACCCACCTCACATGGGCTGTTCTGTTCCTATAGCCTTAAAAA 7333  
Qy 5102 GAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATAGCAAAATGCTGTATGGAGGC 5161  
Db 7334 GAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATAGCAAAATGCTGTATGGAGGC 7393  
Qy 5162 CTTTCAATCAATGACCTTTGTGTGACCCCAAGACAGCCAACTTAGTTGCAGACATCAC 5221  
Db 7394 CTTTCAATCAATGACCTTTGTGTGACCCCAAGACAGCCAACTTAGTTGCAGACATCAC 7453  
Qy 5222 CTCCTTAGCCAAATATCAAGTTCTTAAAAACATTAAGAGAACTTATCCCTGAGAGAA 5281  
Db 7454 CTCCTTAGCCAAATATCAAGTTCTTAAAAACATTAAGAGAACTTATCCCTGAGAGAA 7513  
Qy 5282 GGGAAAGAACTATTCACCCCMWGTGACATGTAATAGTCAAGTCCCTTCCTTAATTC 5341  
Db 7514 GGGAAAGAACTATTCACCCCMWGTGACATGTAATAGTCAAGTCCCTTCCTTAATTC 7573  
Qy 5342 CCCATCCCTAGATACATCTCTGGGAAGGACCTTACCCAGTCAATTTATATACCCCACTGC 5401  
Db 7574 CCCATCCCTAGATACATCTCTGGGAAGGACCTTACCCAGTCAATTTATATACCCCACTGC 7633  
Qy 5402 GGTAAAGTGGCTGGAGTGGAGTCTTGATATACATCACTTTAGTCAAACTCTGGATAC 5461  
Db 7634 GGTAAAGTGGCTGGAGTGGAGTCTTGATATACATCACTTTAGTCAAACTCTGGATAC 7693  
Qy 5462 GCCAAGGAACTGAAAATCCAGAGAACACGCTAGCTATCTCTGTGAACCTCTAGAGGA 5521  
Db 7694 GCCAAGGAACTGAAAATCCAGAGAACACGCTAGCTATCTCTGTGAACCTCTAGAGGA 7753  
Qy 5522 TTTCGCGCTCTCTCAAAACAAACAGAGAGAAAGTAACTAAATCAATAATCCCCCA 5581  
Db 7754 TTTCGCGCTCTCTCAAAACAAACAGAGAGAAAGTAACTAAATCAATAATCCCCCA 7812  
Qy 5582 TGGSCCTCCCTTATCATATTTTCTCTKTASTGTSTTTVACCCCTSTTTTCACTCTCACTG 5641  
Db 7813 TGGSCCTCCCTTATCATATTTTCTCTTACTGTCTTTTACCTCTTTCACCTCTCACTG 7872  
Qy 5642 CACCCCTCATGCGCGTGTATGACAGTAGTCTCCCTVACCMAGAGTTTCTATGAGAA 5701  
Db 7873 CACCCCTCATGCGCGTGTATGACAGTAGTCTCCCTTACCAAGATTTCTATGAGAA 7932  
Qy 5702 TGCAGAGTCCCGGAAATATGATGCCCATCGTATAGGAGTCTTTSTAGGGAACCCCA 5761  
Db 7933 TGCAGAGTCCCGGAAATATGATGCCCATCGTATAGGAGTCTTTTAAAGGGAACCCCA 7992  
Qy 5762 CTTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTTGGCCACTCTTTTGCATGC 5821  
Db 7993 CTTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTTGGCCACTCTTTTGCATGC 8052  
Qy 5822 ATGCAAAATCTATTTGACAGGAAAATGATTAATCTAGTGTCTCTGGAGGACTTG 5881  
Db 8053 ATGCAAAATCTATTTGACAGGAAAATGATTAATCTAGTGTCTCTGGAGGACTTG 8112  
Qy 5882 GAGTCACTGCTCTTGGACTTACTTACCCAACTGATGCTCTGATGGGGTGGAGTTC 5941  
Db 8113 GAGTCACTGCTCTTGGACTTACTTACCCAACTGATGCTCTGATGGGGTGGAGTTC 8172  
Qy 5942 AAGATCAGGCAAGAGAAAACATGTAAAGAACTAATCTCCCAACTCACCGGGTACATG 6001  
Db 8173 AAGATCAGGCAAGAGAAAACATGTAAAGAACTAATCTCCCAACTCACCGGGTACATG 8232  
Qy 6002 GCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAAATACATGAACCCCTCCGTACCC 6061  
Db 8233 GCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAAATACATGAACCCCTCCGTACCC 8292  
Qy 6062 ATACTGCGCTGGTAAGCTATTTTAATACCACTCTCATGCGGCTCCATGAGGCTCGGCC 6121

Db 8293 ATACTGCGCTGGTAAGCTATTTTAATACCACTCCTCACTGGGCTCCATGAGGCTCTGGCCCC 8352  
Qy 6122 AAAACCCCTACTAACTGTTGGATATGCTCCCTGAACTTTCAAGCCCATATGTTTCAATCC 6181  
Db 8353 AAAACCCCTACTAACTGTTGGATATGCTCCCTGAACTTTCAAGCCCATATGTTTCAATCC 8412  
Qy 6182 CTGTACTGCAACAACTGGAACACTTCAGCAGCAAAATAAACACACTTCCTCGTTTAGTAG 6241  
Db 8413 CTGTACTGCAACAACTGGAACACTTCAGCAGCAAAATAAACACACTTCCTCGTTTAGTAG 8472  
Qy 6242 GACCTCTGTTTCCAAATGGAATAAACCATACCTCAAACTCCTGCTGTGTAATAATTA 6301  
Db 8473 GACCTCTGTTTCCAAATGGAATAAACCATACCTCAAACTCCTGCTGTGTAATAATTA 8532  
Qy 6302 GCAATCTAATATACAAACAACTCCCAATGCAATGAGTGGGTAATCTCTCCACACAAA 6361  
Db 8533 GCAATCTAATATACAAACAACTCCCAATGCAATGAGTGGGTAATCTCTCCACACAAA 8592  
Qy 6362 TAGTCTGCGCTTACCTCAGGAATATTTTGTCTGTGTACTTACCTCAGGCTATCGTTGTTGA 6421  
Db 8593 TAGTCTGCGCTTACCTCAGGAATATTTTGTCTGTGTACTTACCTCAGGCTATCGTTGTTGA 8652  
Qy 6422 ATGCGCTCTTCAGAACTCTATGTGCTTCTCTCATTTCTTAGTGGCCCVATGRCATCTACA 6481  
Db 8653 ATGCGCTCTTCAGAACTCTATGTGCTTCTCTCATTTCTTAGTGGCCCTATGACCATCTACA 8712  
Qy 6482 CTGAACAAGATTTATACAGTTATGTCTATCTAAGCCCGCAACAAAGAGTAGTACCCATTC 6541  
Db 8713 CTGAACAAGATTTATACAGTTATGTCTATCTAAGCCCGCAACAAAGAGTAGTACCCATTC 8772  
Qy 6542 TTCTTTTGTATAGAGCAGAGTCTAGTGTGCACTAGTACTTGGCATTTGGCGGTATCA 6601  
Db 8773 TTCTTTTGTATAGAGCAGAGTCTAGTGTGCACTAGTACTTGGCATTTGGCGGTATCA 8832  
Qy 6602 CAACCTCTACTGTTCTACTACAACTATCTCAAGAACTTAATGCGGACATGGAACGG 6661  
Db 8833 CAACCTCTACTGTTCTACTACAACTATCTCAAGAACTTAATGCGGACATGGAACGG 8892  
Qy 6662 TGCGCCACTCCCTGGTCACTTGGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTTC 6721  
Db 8893 TGCGCCACTCCCTGGTCACTTGGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTTC 8952

## RESULT 16

BD221808  
LOCUS BD221808 10499 bp DNA linear PAT 17-JUL-2003  
DEFINITION Nucleic sequence and deduced protein sequence family with human endogenous retroviral motifs, and their uses.  
ACCESSION BD221808  
VERSION BD221808.1 GI:33031578  
KEYWORDS JP 2002518051-A/3.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 10499)  
AUTHORS Alliel, P. M., Perin, J. P. and Rieger, F.  
TITLE Nucleic sequence and deduced protein sequence family with human endogenous retroviral motifs, and their uses  
JOURNAL Patent: JP 2002518051-A 3 25-JUN-2002;  
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE  
COMMENT OS Homo sapiens (human)  
PN JP 2002518051-A/3  
PD 25-JUN-2002  
PF 23-JUN-1999 JP 2000556036  
PR 23-JUN-1998 FR 98/07920  
PI PATRICK M ALLIEL, JEAN PIERRE PERIN, FRANCOIS RIEGER PC  
C12N15/09, A01K67/027, A61K31/713, A61K39/21, A61K48/00, A61P21/00, PC  
A61P25/00,  
PC A61P37/06, C07K14/15, C12Q1/68, C12Q1/70, C12N15/00 CC  
sequence and deduced protein sequence family with CC  
human endogenous

CC		retroviral motifs, and their uses	
Key	Location/Qualifiers		
FT	1..10499		
FT	Location/Qualifiers		
source	1..10499		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	10.3%;	Score 779;	DB 2; Length 10499;
Best Local Similarity	98.9%;	Pred. No. 0;	
Matches 1899;	Conservative 0;	Mismatches 20;	Indels 1; Gaps 1;
QY	4802	AGGCTTCCCTCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAAGGCACTAGTTTCATGAAA	4861
Db	7101	AGGCTTCCCTCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAAGGCACTAGTTTCATGAAA	7160
QY	4862	TAAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCACGG	4921
Db	7161	TAAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCACGG	7220
QY	4922	CCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTCGCGCT	4981
Db	7221	CCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTCGCGCT	7280
QY	4982	GAAAGCCACAGTCTCTCAGGGAAGGTCCAGAAAATGAATGAAAYACTCAAAGGACATCTAA	5041
Db	7281	GAAAGCCACAGTCTCTCAGGGAAGGTCCAGAAAATGAATGAAAYACTCAAAGGACATCTAA	7340
QY	5042	AAAAAGCAACCCAGGAAACCCACCTCACATGGCCTGTCTCTGTGCCCTATAGCCTTAAAAA	5101
Db	7341	AAAAAGCAACCCAGGAAACCCACCTCACATGGCCTGTCTCTGTGCCCTATAGCCTTAAAAA	7400
QY	5102	GAACTGTGAACCTTTCCCAAAAGAGGAGCTTAGGCCATACGAAATGCTGTATGGAAGGC	5161
Db	7401	GAACTGTGAACCTTTCCCAAAAGAGGAGCTTAGGCCATACGAAATGCTGTATGGAAGGC	7460
QY	5162	CCTTCATAACCAATGACCTTGTGCTTGACCCAGACAGCCAACTTAGTTGCAGACATCAC	5221
Db	7461	CCTTCATAACCAATGACCTTGTGCTTGACCCAGACAGCCAACTTAGTTGCAGACATCAC	7520
QY	5222	CTCCTTAGCCAAATATCAACAAAGTCTTTAAACATTTACAGGAACTTACCTCGAGAAGA	5281
Db	7521	CTCCTTAGCCAAATATCAACAAAGTCTTTAAACATTTACAGGAACTTACCTCGAGAAGA	7580
QY	5282	GGGAAAAGAACTATTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAAATTC	5341
Db	7581	GGGAAAAGAACTATTCCACCCCTTGTGACATGGTATTAGTCAAGTCCCTTCCTCTAAATTC	7640
QY	5342	CCCATCCCTAGATACATCTCTGGGAAGGACCCCTACCCAGTCACTTTATYTTACCCCACTGC	5401
Db	7641	CCCATCCCTAGATACATCTCTGGGAAGGACCCCTACCCAGTCACTTTTATCTACCCCACTGC	7700
QY	5402	GGTTAAAGTGGCTGGAGTGGAGTCTTGATACATCACACTTGAATCAATCTCGGATACT	5461
Db	7701	GGTTAAAGTGGCTGGAGTGGAGTCTTGATACATCACACTTGAATCAATCTCGGATACT	7760
QY	5462	GCCAAAGGAACTGAAAATCCAGGAGCAACCGCTAGCTATTCTCTGTGAACCTCTAGAGGA	5521
Db	7761	GCCAAAGGAACTGAAAATCCAGGAGCAACCGCTAGCTATTCTCTGTGAACCTCTAGAGGA	7820
QY	5522	TTTGGCGCTGCTCTTCAAAACAAACACAGGAGGAAAGTAACATAAATCATAAATCCCCCA	5581
Db	7821	TTTGGCGCTGCTCTTCAAAACAAACACAGGAGGAAAGTAACATAAATCATAAATCCCCCA	7879
QY	5582	TGSGCTCCCTTATCATATTTTCTCTKTAGTGTSTTTTACCCTSTTTTCACTCTCACTG	5641
Db	7880	TGSGCTCCCTTATCATATTTTCTCTTACTGTCTTTTACCCTCTTTTCACTCTCACTG	7939
QY	5642	CACCCCTCCATGCCGCTGTATGACCAGTAGCTCCCTTACCMAGAGTTTCTATGGAGAA	5701

DB	7940	CACCCCTCCATGCGCTGTATGACACAGTAGCTCCCTTTACCAAGAGTTTCTATGAGAA	7999				
QY	5702	TGCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTSTTAAGGAAACCCCA	5761				
DB	8000	TGCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTCTAAGGAAACCCCA	8059				
QY	5762	CCTTCACCTGCCACACCCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTTCATGC	5821				
DB	8060	CCTTCACCTGCCACACCCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTTCATGC	8119				
QY	5822	ATGCAATATCATATTATTGGACAGAAAATGATTAATCTTAGTTGTCTCTGGAGACTTG	5881				
DB	8120	ATGCAATATCATATTATTGGACAGAAAATGATTAATCTTAGTTGTCTCTGGAGACTTG	8179				
QY	5882	GAGTCACCTGTCTGTGGACTTACTTCAACCAACTGGTATGCTCTATGGGGTGGAGTTC	5941				
DB	8180	GAGTCACCTGTCTGTGGACTTACTTCAACCAACTGGTATGCTCTATGGGGTGGAGTTC	8239				
QY	5942	AAGATCAGGCAAGAGAAAACATGTAAAAGAAATCTCTCCCAACTCACCCGGGTACATG	6001				
DB	8240	AAGATCAGGCAAGAGAAAACATGTAAAAGAAATCTCTCCCAACTCACCCGGGTACATG	8299				
QY	6002	GCACTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCCTCCGTACCC	6061				
DB	8300	GCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCCTCCGTACCC	8359				
QY	6062	ATACTCGGCTGGTAAGCCTATTTAATACCACTCCTCAGTGGGCTCATGAGTCTCGGCC	6121				
DB	8360	ATACTCGGCTGGTAAGCCTATTTAATACCACTCCTCAGTGGGCTCATGAGTCTCGGCC	8419				
QY	6122	AAAAACCTTAACTGTTGGATATGCTCCCTCGAACTTCARGGCATATGTTTCAATCC	6181				
DB	8420	AAAAACCTTAACTGTTGGATATGCTCCCTCGAACTTCARGGCATATGTTTCAATCC	8479				
QY	6182	CTGTACTCTGAAACAAATTTGAAACAACTTTCAGCAGAAATAAACCACTTCCGTTTTAGTAG	6241				
DB	8480	CTGTACTCTGAAACAAATTTGAAACAACTTTCAGCAGAAATAAACCACTTCCGTTTTAGTAG	8539				
QY	6242	GACCTCTGTTTCCNAATSTGGNAATAACCCATACCTCAAACTCAGCTGTGTAAATTTA	6301				
DB	8540	GACCTCTGTTTCCNAATSTGGNAATAACCCATACCTCAAACTCAGCTGTGTAAATTTA	8599				
QY	6302	GCAATCTACATACACAAACCAACTCCCAATGATCAGGTGGGTAACTCTCTCCACACAAA	6361				
DB	8600	GCAATCTACATACACAAACCAACTCCCAATGATCAGGTGGGTAACTCTCTCCACACAAA	8659				
QY	6362	TAGTCTGCCTACCTCAGGAAATATTTTGTCTGTGGTACCTCAGCCTATCGTTGTGA	6421				
DB	8660	TAGTCTGCCTACCTCAGGAAATATTTTGTCTGTGGTACCTCAGCCTATCGTTGTGA	8719				
QY	6422	ATGGCTCTTCAAGATCTATGCTTCTCTCATTTCTTAGTGGCCCCYATGRCATCTACA	6481				
DB	8720	ATGGCTCTTCAAGATCTATGCTTCTCTCATTTCTTAGTGGCCCCYATGRCATCTACA	8779				
QY	6482	CTGAAACAGATTTATACAGTTATGTCATATCTTAAGCCCGCAACAAAAGAGTACCCATTC	6541				
DB	8780	CTGAAACAGATTTATACAGTTATGTCATATCTTAAGCCCGCAACAAAAGAGTACCCATTC	8839				
QY	6542	TTCTTTTGTATATAGGACGAGAGTGTAGGTGCACTAGGTACTGGCAATGGCGGTATCA	6601				
DB	8840	TTCTTTTGTATATAGGACGAGAGTGTAGGTGCACTAGGTACTGGCAATGGCGGTATCA	8899				
QY	6602	CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGACGGG	6661				
DB	8900	CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGACGGG	8959				
QY	6662	TGCGCGACTCCCTGTCTGACCTTGGCAAGATCAACTTAACTCCCTAGCAGAGTAGTCTCTTC	6721				
DB	8960	TGCGCGACTCCCTGTCTGACCTTGGCAAGATCAACTTAACTCCCTAGCAGAGTAGTCTCTTC	9019				



LOCUS	AR699757	10499 bp	DNA	linear	PAT 14-SEP-2000
DEFINITION	Sequence 3 from patent US 6919438.				
ACCESSION	AR699757				
VERSION	AR699757.1	GI:75205546			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 10499)				
TITLE	Alliel,P.M., Perin,J.-P. and Rieger,F.				
JOURNAL	Nucleic sequence and deduced protein sequence family with human endogenous retroviral motifs, and their uses				
	Patent: US 6919438-A 3 19-JUL-2005;				
	Institut National de la Sante et de la Recherche Medical-Inserm;				
	Paris;				
	FRX;				
FEATURES	Location/Qualifiers				
source	1..10499				
	/organism="unknown"				
ORIGIN	/mol_type="genomic DNA"				
	Query Match	10.3%;	Score 779;	DB 2;	Length 10499;
	Best Local Similarity	98.9%;	Pred. No. 0;		
	Matches 1899;	Conservative	0;	Mismatches	20;
				Indels	1;
				Gaps	1;
Qy	4802	AGGCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAAGGCACATAGTCTTCATGAAA	4861		
Db	7101	AGGCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAAGGCACATAGTCTTCATGAAA	7160		
Qy	4862	TAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACATAGCCCTGCTTTCGAGG	4921		
Db	7161	TAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACATAGCCCTGCTTTCGAGG	7220		
Qy	4922	CCAAGTAAACCCAGGAGTATCCAGGCGTTTGGTATACGATATCACTTACACTGCGCCT	4981		
Db	7221	CCAAGTAAACCCAGGAGTATCCAGGCGTTTGGTATACGATATCACTTACACTGCGCCT	7280		
Qy	4982	GAAGGCCACAGTCCTCAGGGAAGGTTCAGAAAAATGAATGAAAYACTCAAAAGGACATCTAA	5041		
Db	7281	GAAGGCCACAGTCCTCAGGGAAGGTTCAGAAAAATGAATGAAAYACTCAAAAGGACATCTAA	7340		
Qy	5042	AAAAGCAACCCAGGAAAACCCACTCATATGGGCTGTTCTGTTGGCTATATAGCCTTAAAAA	5101		
Db	7341	AAAAGCAACCCAGGAAAACCCACTCATATGGGCTGTTCTGTTGGCTATATAGCCTTAAAAA	7400		
Qy	5102	GAATCTGCAACTTTCCTCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC	5161		
Db	7401	GAATCTGCAACTTTCCTCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC	7460		
Qy	5162	CCCTTCAATACCAATGACCTTGTCTTGACCCCAAGACAGCCAACTTAGTTGCGACACATCAC	5221		
Db	7461	CCCTTCAATACCAATGACCTTGTCTTGACCCCAAGACAGCCAACTTAGTTGCGACACATCAC	7520		
Qy	5222	CTCCTTAGCCAAATATCAACAAGTCTTTAAAAATTACAAGGAACCTATCCCTCGAAGA	5281		
Db	7521	CTCCTTAGCCAAATATCAACAAGTCTTTAAAAATTACAAGGAACCTATCCCTCGAAGA	7580		
Qy	5282	GGGAAAGAACTATTCCACCCWGTGACATGTTAGTCAGTCCCTTCYCTCTAATTC	5341		
Db	7581	GGGAAAGAACTATTCCACCCWGTGACATGTTAGTCAGTCCCTTCYCTCTAATTC	7640		
Qy	5342	CCCATCCCTAGATACATCTCTGGGAAGGACCTTACCAGTCACTTTTATATACCCCACTGC	5401		
Db	7641	CCCATCCCTAGATACATCTCTGGGAAGGACCTTACCAGTCACTTTTATATACCCCACTGC	7700		
Qy	5402	GGTTAAAGTGGCTGGAGTGGAGTCTTTGGATACATCACTTTGAGTCAAACTCTGTGATACT	5461		
Db	7701	GGTTAAAGTGGCTGGAGTGGAGTCTTTGGATACATCACTTTGAGTCAAACTCTGTGATACT	7760		
Qy	5462	GCCAAAGGAACCTGAAAAATCCAGGAGACAAACGCTAGCTATTCTCTGTGAACCTCTAGAGA	5521		
Db	7761	GCCAAAGGAACCTGAAAAATCCAGGAGACAAACGCTAGCTATTCTCTGTGAACCTCTAGAGA	7820		

Qy	6602	CAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTTAATGGGGACATGGAACGGG	6666
Db	8900	CAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTTAATGGGGACATGGAACGGG	8959
Qy	6662	TCGCCGAGCTCCCTGGTCACCTTCGAAAGTCAAACTTAACTCCCTAGCAGCAGTAGTCTCTTC	6721
Db	8960	TCGCCGAGCTCCCTGGTCACCTTCGAAAGTCAAACTTAACTCCCTAGCAGCAGTAGTCTCTTC	9019
RESULT 18			
AX007980		LOCUS	
DEFINITION		Sequence 3 from Patent WO9967395.	
ACCESSION		AX007980	
VERSION		AX007980.1 GI:9995677	
KEYWORDS		Homo sapiens (human)	
SOURCE		Homo sapiens	
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae; Homo.	
REFERENCE		1	
AUTHORS		Perin,J.P., Rieger,F. and Alliel,P.M.	
TITLE		Nucleic sequence and deduced protein sequence family with human endogenous retroviral motifs, and their uses	
JOURNAL		Patent: WO 9967395-A 3 29-DEC-1999; INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER FRANCOIS (FR); ALLIEL PATRICK M (FR)	
FEATURES		Location/Qualifiers	
source		1..10499	
ORIGIN		/organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"	
Query Match 10.3%; Score 779; DB 2; Length 10499;			
Best Local Similarity 98.9%; Pred. No. 0;			
Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;			
Qy	4802	AGGCCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCATGAAA	4861
Db	7101	AGGCCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCATGAAA	7160
Qy	4862	TAATTCACAGATTCGGAGCTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCAGG	4921
Db	7161	TAATTCACAGATTCGGAGCTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCAGG	7220
Qy	4922	CCACAGTAACCCAGGAGTAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCT	4981
Db	7221	CCACAGTAACCCAGGAGTAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCT	7280
Qy	4982	GAAGGCACAGTCTCTAGGGAGGTCCAGAAAATGAATGAAYAACTCAAAAGGACATCTAA	5041
Db	7281	GAAGGCACAGTCTCTAGGGAGGTCCAGAAAATGAATGAAYAACTCAAAAGGACATCTAA	7340
Qy	5042	AAAGAGCAACCCAGGAAACCCACCTACATGGCCGTCTGTTCGCTATAGCCCTTAAAAA	5101
Db	7341	AAAGAGCAACCCAGGAAACCCACCTACATGGCCGTCTGTTCGCTATAGCCCTTAAAAA	7400
Qy	5102	GAATCTGCAACTTTTCCCAAAAAAGCAGGACTTAGCCCATAGAAATGCTGTATGGAAAGGC	5161
Db	7401	GAATCTGCAACTTTTCCCAAAAAAGCAGGACTTAGCCCATAGAAATGCTGTATGGAAAGGC	7460
Qy	5162	CTTTCATAAACAATGACCTTGTGCTTGACCCAGACAGCCAACTTAGTTGGAGACATCAC	5221
Db	7461	CTTTCATAAACAATGACCTTGTGCTTGACCCAGACAGCCAACTTAGTTGGAGACATCAC	7520
Qy	5222	CTCCTTAGCCAAATATCAACAAGTTCTTAAACATTACAAGAACCTTATCCCTGAGAAGA	5281
Db	7521	CTCCTTAGCCAAATATCAACAAGTTCTTAAACATTACAAGAACCTTATCCCTGAGAAGA	7580
Qy	5282	GGGAAAAAGAACTATTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTC	5341



Qy	6422	ATGGCTCTTCAGAACTCTATGTGCTTCTCTCATCTTTAGTGCCCCCVATGRCATCTACA	6481
Db	8720	ATGGCTCTTCAGAACTCTATGTGCTTCTCTCATCTTTAGTGCCCCCVATGRCATCTACA	8779
Qy	6482	CTGAACAAGATTTATACAGTTATGTCTATATCTTAAGCCCCGCAACAAAGAAGTAGTACCCATTTC	6541
Db	8780	CTGAACAAGATTTATACAGTTATGTCTATATCTTAAGCCCCGCAACAAAGAAGTAGTACCCATTTC	8839
Qy	6542	TTCCCTTTTGTATTAGGAGCAGGAGTGCTAGGTGCACCTAGGTACTTGGCATTTGGCGGTATCA	6601
Db	8840	TTCCCTTTTGTATTAGGAGCAGGAGTGCTAGGTGCACCTAGGTACTTGGCATTTGGCGGTATCA	8899
Qy	6602	CAACCTCTACTCAGTTCTTACTACAAACTATCTCAAGAACTAAATGGGGACATCGAACGGG	6661
Db	8900	CAACCTCTACTCAGTTCTTACTACAAACTATCTCAAGAACTAAATGGGGACATCGAACGGG	8959
Qy	6662	TCGCCGACTCCCTGGTCACCTTGCAGAGTCAACTTAACCTCCCTAGCAGCAGTAGTCTCTTC	6721
Db	8960	TCGCCGACTCCCTGGTCACCTTGCAGAGTCAACTTAACCTCCCTAGCAGCAGTAGTCTCTTC	9019
RESULT 19			
LOCUS	AX329572 56093 bp DNA linear PAT 09-JAN-2002		
DEFINITION	Sequence 81 from Patent WO0194629.		
ACCESSION	AX329572		
VERSION	AX329572.1 GI:18102550		
KEYWORDS	.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Homnidae; Homo.		
	1		
AUTHORS	Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,		
	Horrigan, S., Soppet, D.R. and Weaver, Z.		
TITLE	Cancer gene determination and therapeutic screening using signature		
JOURNAL	gene sets		
FEATURES	Patent: WO 0194629-A 81 13-DEC-2001;		
	Avalon Pharmaceuticals (US)		
	Location/Qualifiers		
ORIGIN	1..56093		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
/db_xref="taxon:9606"			
Query Match 10.13%; Score 779; DB 2; Length 56093;			
Best Local Similarity 98.9%; Pred. No. 0;			
Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;			
Qy	4802	AGGCCTTCCCTCTAGGACAGAAAAGCCCAAGAGGTAAATAAGGCACCTAGTTCATGAAA	4861
Db	35101	AGGCCTTCCCTCTAGGACAGAAAAGCCCAAGAGGTAAATAAGGCACCTAGTTCATGAAA	35160
Qy	4862	TAATTCCAGATTCCGACTTCCCGAGGCTTACAGAGTGCACATAGCCCTGCTTTCCAGG	4921
Db	35161	TAATTCCAGATTCCGACTTCCCGAGGCTTACAGAGTGCACATAGCCCTGCTTTCCAGG	35220
Qy	4922	CCACAGTAACCCAGGAGTAGTATCCAGCGGTAGGTATACGATATCATCTACATGCGCCT	4981
Db	35221	CCACAGTAACCCAGGAGTAGTATCCAGCGGTAGGTATACGATATCATCTACATGCGCCT	35280
Qy	4982	GAAGGCCACACAGTCTCTAGGGAAGGTGCAGAAAATGAATGAAYACTCTCAAGGACATCTAA	5041
Db	35281	GAAGGCCACACAGTCTCTAGGGAAGGTGCAGAAAATGAATGAAYACTCTCAAGGACATCTAA	35340
Qy	5042	AAAAGCAAAACCCAGGAAAACCCACCTCACATGGCGCTGTCCTGTTGCCCTATAGCCTTAAAAA	5101
Db	35341	AAAAGCAAAACCCAGGAAAACCCACCTCACATGGCGCTGTCCTGTTGCCCTATAGCCTTAAAAA	35400
Qy	5102	GAATCTGCACACTTTCCTCCCAAAAAGCAGGACTTAGGCCCATACGAAAATGCTGTATGGAAGGC	5161



```

/rpt family="ALU"
13878..13906
/genes="WUGSC:H.RG083M05.1"
/notes="match to human 5' EST H41382 (NID:g917434), bases
30-58"
repeat_region
13907..14104
/rpt family="ALU"
complement(14110..14137)
/rpt family="L1"
repeat_region
complement(15618..15907)
/rpt family="ALU"
17227..17522
/rpt family="ALU"
18667..19235
/notes="match to human fetal brain 5' EST D61494
(NID:g970409), bases 1-255, and to human 3' EST R07476
(NID:g759399)"
repeat_region
19550..19670
/rpt family="ALU"
21507..37303
/notes="similarity to various SS-RNA virus polyproteins;
pseudogene; region of matches and close matches to
multiple human ESTs, see R68740 (NID:g842257)"
exon
37316..37489
/inference="non-experimental evidence, no additional
details recorded"
/notes="Grail prediction, score = 80"
repeat_region
complement(38938..39224)
/rpt family="ALU"
39225..39707
/notes="match to multiple human ESTs, see N30113
(NID:g1148633)"
repeat_region
39800..40085
/rpt family="ALU"
complement(40247..40538)
/rpt family="ALU"
repeat_region
complement(40632..40924)
/rpt family="ALU"
repeat_region
complement(42283..42891)
/rpt family="ALU"
repeat_region
complement(45474..45613)
/rpt family="ALU"
repeat_region
complement(45614..45737)
/notes="match to human 3' EST H48898 (NID:g988738), bases
129-333"
misc_feature
complement(46107..47026)
/notes="match to multiple human ESTs, see N81064
(NID:g1243765), H48897 (NID:g988737), and M78831
(NID:g273146)"
repeat_region
complement(47027..47318)
/rpt family="ALU"
misc_feature
complement(47365..47782)
/notes="match to multiple human ESTs, see W37495
(NID:g1319089)"
misc_feature
47898..48115
/notes="match to human 5' EST H62306 (NID:g1015138), bases
93-368"
repeat_region
complement(48116..48405)
/rpt family="ALU"
misc_feature
complement(48406..48584)
/notes="match to human 3' EST N29952 (NID:g1148472), bases
290-455, and 5' EST R12730 (NID:g765806)"
repeat_region
complement(48787..49405)
/rpt family="ALU"
misc_feature
complement(49406..49534)
/notes="match to human 3' EST R65794 (NID:g838432), bases
309-440"
repeat_region
complement(49638..49672)
/rpt family="ALU"
misc_feature
complement(49674..49890)
/notes="match to human 3' EST N29952 (NID:g1148472) and 5'
EST N29938 (NID:g1148458), sequences are from opposite
ends of the same clone"

```

---

```

gene
complement(49698..51806)
/genes="WUGSC:H.RG083M05.2"
complement(join(49698..49888,51575..51806))
CDS
/genes="WUGSC:H.RG083M05.2"
/notes="coded for by human cDNAs W37389 (NID:g1319205),
R65891 (NID:g838529), R65794 (NID:g838432) and R65794
(NID:g838432)"
/codon_start=1
/protein_id="AAB46345.1"
/db_xref="GI:1669370"
/translation="MEFYFQCGIIFCFPGQVYQIGDVVSVIDEODGKPYAQIRGF
IDQVCEKSAALTWLIPTLSSPDQDPASVLIIGPREDLPKMEYLEFVCHAFSEYFK
SRSSPPTVPTPEKGYIWHVGPPTAIKESVANHL"
exon
complement(51576..51758)
/genes="WUGSC:H.RG083M05.2"
/inference="non-experimental evidence, no additional
details recorded"
/notes="Grail prediction, score = 86"

```

---

```

Query Match      10.3%; Score 779; DB 5; Length 56093;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 4802 AGGCTTCCCTCTAGGACAGAAAAGCCCAAGAGGTATAAAGGCACTAGTTTCATGAAA 4861
      |||||
Db 35101 AGGCTTCCCTCTAGGACAGAAAAGCCCAAGAGGTATAAAGGCACTAGTTTCATGAAA 35160

QY 4862 TAATTCGAGATTGCGACTTCCCGAGAGCTTACAGAGTGAACAATAGCCCTGCTTCCAGG 4921
      |||||
Db 35161 TAATTCGAGATTGCGACTTCCCGAGAGCTTACAGAGTGAACAATAGCCCTGCTTCCAGG 35220

QY 4922 CCACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCT 4981
      |||||
Db 35221 CCACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCT 35280

QY 4982 GRAGGCCACAGTCTCAGGGAAGGTCGAGAAATGAATGAATACTCAAGGACATCTAA 5041
      |||||
Db 35281 GRAGGCCACAGTCTCAGGGAAGGTCGAGAAATGAATGAATACTCAAGGACATCTAA 35340

QY 5042 AAAAGCAAAACCCAGGAAAACCCACTCACATGGCTGCTGTGTTGCCCTATAGCCCTTAAAAA 5101
      |||||
Db 35341 AAAAGCAAAACCCAGGAAAACCCACTCACATGGCTGCTGTGTTGCCCTATAGCCCTTAAAAA 35400

QY 5102 GAATCTGCAACTTTCCTCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 5161
      |||||
Db 35401 GAATCTGCAACTTTCCTCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 35460

QY 5162 CTTTCATACCAATGACCTTGTGTTGACCAAGACAGCCAACTTAGTTAGTCAGACATCAC 5221
      |||||
Db 35461 CTTTCATACCAATGACCTTGTGTTGACCAAGACAGCCAACTTAGTTAGTCAGACATCAC 35520

QY 5222 CTCTTTAGCCAAATATCAACAAGTTCTTAAACATTTACAAGGAACCTTATCCCTGAGAGA 5281
      |||||
Db 35521 CTCTTTAGCCAAATATCAACAAGTTCTTAAACATTTACAAGGAACCTTATCCCTGAGAGA 35580

QY 5282 GGGAAAAGAACTATTCCACCCWMTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTC 5341
      |||||
Db 35581 GGGAAAAGAACTATTCCACCCCTTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTC 35640

QY 5342 CCCATCCCTAGATACATCTCTGGGAAGGACCTTACCCAGTCATTTTATYTAACCCAACTGC 5401
      |||||
Db 35641 CCCATCCCTAGATACATCTCTGGGAAGGACCTTACCCAGTCATTTTATYTAACCCAACTGC 35700

QY 5402 GGTTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACTTGGAGTCAAAATCTGGGACT 5461
      |||||
Db 35701 GGTTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACTTGGAGTCAAAATCTGGGACT 35760

QY 5462 GCCAAAGAACTGAAAATCCAGGAGACAAACGCTAGCTATTCTGTGAACCTCTAGAGGA 5521
      |||||
Db 35761 GCCAAAGAACTGAAAATCCAGGAGACAAACGCTAGCTATTCTGTGAACCTCTAGAGGA 35820

QY 5522 TTTTGGCGCTGCTTTTCAAAACAAACACAGGAGAAAGTAACTAAATCATAAATCCCCCA 5581
      |||||

```

Db 35821 TTTGGCCTGCTCTTCAACACACACACAGGAGGAAAGTAACTAAATCATAAAT-CCCCA 35879  
Qy 5582 TGGSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCCTSTTTTCACTCTCACTG 5641  
Db 35880 TGGCCCTCCCTTATCATATTTTCTCTTTTACTGTTCTTTTACCCTCTTTCACTCTCACTG 35939  
Qy 5642 CACCCCTCCATGCGCGTGTATGACAGTAGCTCCCTTACCMAGAGTTTCTATGGAGAA 5701  
Db 35940 CACCCCTCCATGCGCGTGTATGACAGTAGCTCCCTTACCAAGATTTCTATGGAGAA 35999  
Qy 5702 TGCAGCGTCCCGAAATATGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCA 5761  
Db 36000 TGCAGCGTCCCGAAATATGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCA 36059  
Qy 5762 CTTTCACTGCCACACCCATATGCGCGCAACTGCTATCACTCTGCCACTCTTTGCATGC 5821  
Db 36060 CTTTCACTGCCACACCCATATGCGCGCAACTGCTATCACTCTGCCACTCTTTGCATGC 36119  
Qy 5822 ATGCAAAATCACTATTATGGACAGGAAATGATTAATCCTAGTTGTCTCTGGAGACTTG 5881  
Db 36120 ATGCAAAATCACTATTATGGACAGGAAATGATTAATCCTAGTTGTCTCTGGAGACTTG 36179  
Qy 5882 GAGTCACTCTGTGTGGACTTACTTCAACCAACTGGTATGCTGATGGGGTGGAGTTC 5941  
Db 36180 GAGTCACTCTGTGTGGACTTACTTCAACCAACTGGTATGCTGATGGGGTGGAGTTC 36239  
Qy 5942 AAGATCAGGCAAGAGAAAACATGTAAAGAGTATCTCCCACTCACCCGGGTACATG 6001  
Db 36240 AAGATCAGGCAAGAGAAAACATGTAAAGAGTATCTCCCACTCACCCGGGTACATG 36299  
Qy 6002 GCACCTCTAGCCCTACAAAGACATGATCTCTCAAACTACATGAACCCCTCCGTACCC 6061  
Db 36300 GCACCTCTAGCCCTACAAAGACATGATCTCTCAAACTACATGAACCCCTCCGTACCC 36359  
Qy 6062 ATACTCGCTGGTAAAGCTTATTAATACCAACCTCACTGGGCTCCATGAGTCTGGGCC 6121  
Db 36360 ATACTCGCTGGTAAAGCTTATTAATACCAACCTCACTGGGCTCCATGAGTCTGGGCC 36419  
Qy 6122 AAAACCTACTAATCTTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATCC 6181  
Db 36420 AAAACCTACTAATCTTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATCC 36479  
Qy 6182 CTGTACTCAAAATGGAACACTTCAGCACAGAAATAAACCACTCCGTTTGTAGTAG 6241  
Db 36480 CTGTACTCAAAATGGAACACTTCAGCACAGAAATAAACCACTCCGTTTGTAGTAG 36539  
Qy 6242 GACCTCTGTTCCTCAATGGAATAACCATACCTCAAACTCACTGTGTAAATTTA 6301  
Db 36540 GACCTCTGTTCCTCAATGGAATAACCATACCTCAAACTCACTGTGTAAATTTA 36599  
Qy 6302 GCAATACATACACAACTCCCAATGCAATGAGTGGGTAATCTCTCCACACAAA 6361  
Db 36600 GCAATACATACACAACTCCCAATGCAATGAGTGGGTAATCTCTCCACACAAA 36659  
Qy 6362 TAGTCTGCTTACCCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCGTTGTTCA 6421  
Db 36660 TAGTCTGCTTACCCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCGTTGTTCA 36719  
Qy 6422 ATGGCTCTTCAGAACTATGTGTCTCTCTCTCACTTCTAGTGCCCTCCATGACATCTACA 6481  
Db 36720 ATGGCTCTTCAGAACTATGTGTCTCTCTCTCACTTCTAGTGCCCTCCATGACATCTACA 36779  
Qy 6482 CTGAACAAGATTATACAGTTATGTATATCTAAGCCCGCAACAAAGAGTACCCATTC 6541  
Db 36780 CTGAACAAGATTATACAGTTATGTATATCTAAGCCCGCAACAAAGAGTACCCATTC 36839  
Qy 6542 TTCTCTTTTGTATAGGACAGGAGTCTAGTGCATAGTACTGTCATTTGGCGGTATCA 6601  
Db 36840 TTCTCTTTTGTATAGGACAGGAGTCTAGTGCATAGTACTGTCATTTGGCGGTATCA 36899  
Qy 6602 CAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGG 6661  
Db 36900 CAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGG 36959

Qy 6662 TCGCGACTCCCTGCTCACCCTTGAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTC 6721  
Db 36960 TCGCGACTCCCTGCTCACCCTTGAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTC 37019

RESULT 21  
AC007566/c 149194 bp DNA linear PRI 01-MAR-2002  
LOCUS Homo sapiens BAC clone CTB-10G5 from 7q21-7q22, complete sequence.  
DEFINITION AC007566  
ACCESSION AC007566  
VERSION AC007566.2 GI:11181861  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE  
AUTHORS Sulston, J.E. and Waterston, R.  
TITLE Toward a complete human genome sequence  
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)  
PUBMED 9847074

REFERENCE  
AUTHORS Du, Z.  
TITLE The sequence of Homo sapiens BAC clone CTB-10G5  
JOURNAL Unpublished (2001)  
REFERENCE  
AUTHORS Waterston, R.H.  
TITLE Direct Submission  
JOURNAL Submitted (15-MAY-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE  
AUTHORS Waterston, R.  
TITLE Direct Submission  
JOURNAL Submitted (02-OCT-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
REFERENCE  
AUTHORS Waterston, R.H.  
TITLE Direct Submission  
JOURNAL Submitted (16-NOV-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE  
AUTHORS Waterston, R.H.  
TITLE Direct Submission  
JOURNAL Submitted (03-JAN-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
REFERENCE  
AUTHORS Waterston, R.H.  
TITLE Direct Submission  
JOURNAL Submitted (06-FEB-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE  
AUTHORS Waterston, R.  
TITLE Direct Submission  
JOURNAL Submitted (01-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
COMMENT  
On Nov 16, 2000 this sequence version replaced gi:4835815.

----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc>  
Contact: [sapiens@watson.wustl.edu](mailto:sapiens@watson.wustl.edu)  
----- Summary Statistics  
-----  
Center project name: H\_RG010G05

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping

clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send mailto:egreen@nhgri.nih.gov, or see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

Clone CTB-10G5 is from the first release of the human BAC library CTB-978SK-B. The library contains cloned DNA from the male fibroblast cell line 978SK. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8 (1996). This clone is available from Research Genetics, Inc. (<http://www.regen.com>).  
VECTOR: pBelobAC11

Selection: chloramphenicol

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP5-911H5, 200 base pair overlap. Actual start of this clone is at base position 195 of CTB-10G5; actual end is at base position 150532 of CTB-10G5.

The clone CTB-10G5 contains the entire sequence of CTB-83M5.

#### FEATURES

```

source
1..149194
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="7"
    /map="7q21-7q22"
    /clone="CTB-10G5"
    /clone_lib="CTB-978SK-B"
1..1634
    /rpt_family="L1"
repeat_region
2248..2388
    /note="match to EST BG752083 (NID:gl4063536)"
2248..2387
    /note="match to EST AA149693 (NID:gl720635) zn99d06.r1"
2248..2387
    /note="match to EST AW579261 (NID:g7254310)"
2248..2387
    /note="match to EST BG766882 (NID:gl4077535)"
2248..2387
    /note="match to EST BI160365 (NID:gl4620366)"
2248..2374
    /note="similar to Homo sapiens EST BF758865 (NID:gl2106765)"
2248..2287
    /note="match to EST AV686676 (NID:gl0288539)"
2253..2387
    /note="match to EST AU123510 (NID:gl0948226)"
2344..2387
    /note="match to EST AV686676 (NID:gl0288539)"
2696..3066
    /rpt_family="MaLR"
3108..3392
    /rpt_family="Alu"
3340..3628
    /note="match to EST AA425526 (NID:gl2106267) zw48b03.r1"

misc_feature
3540..3628
    /note="similar to Mus musculus EST BB253526 (NID:gb946272)"
3542..3632
    /note="match to EST AV686676 (NID:gl0288539)"
3542..3628
    /note="match to EST AA149693 (NID:gl720635) zn99d06.r1"
3542..3628
    /note="match to EST AU123510 (NID:gl0948226)"
3542..3628
    /note="match to EST BG766882 (NID:gl4077535)"
3542..3628
    /note="match to EST BI160365 (NID:gl4620366)"
3542..3609
    /note="match to EST AW579261 (NID:g7254310)"
3562..3628
    /note="match to EST BE272564 (NID:g9146913)"
3598..3628
    /note="match to EST BE299708 (NID:g9183456)"
3714..3785
    /note="match to EST BG260659 (NID:gl2770475)"
3717..3785
    /note="similar to Mus musculus EST BE994936 (NID:gl0678674)"
3719..3785
    /note="match to EST AA149693 (NID:gl720635) zn99d06.r1"
3719..3785
    /note="match to EST AA425526 (NID:gl2106267) zw48b03.r1"
3719..3785
    /note="match to EST AU123510 (NID:gl0948226)"
3719..3785
    /note="match to EST BE272564 (NID:g9146913)"
3719..3785
    /note="match to EST BE299708 (NID:g9183456)"
3719..3785
    /note="match to EST BG766882 (NID:gl4077535)"
3719..3785
    /note="match to EST BI160365 (NID:gl4620366)"
3719..3785
    /note="similar to Mus musculus EST BB253526 (NID:gb946272)"
3860..3862
    /note="match to EST AV686676 (NID:gl0288539)"
4730..4827
    /rpt_family="L2"
4834..4925
    /rpt_family="L1"
5715..5860
    /note="match to EST AA425526 (NID:gl2106267) zw48b03.r1"
5715..5860
    /note="match to EST AU123510 (NID:gl0948226)"
5715..5860
    /note="match to EST BE272564 (NID:g9146913)"
5715..5860
    /note="match to EST BE299708 (NID:g9183456)"
5715..5860
    /note="match to EST BG260659 (NID:gl2770475)"
5715..5860
    /note="match to EST BG766882 (NID:gl4077535)"
5715..5860
    /note="match to EST BI160365 (NID:gl4620366)"
5715..5860

Query Match
10.3%; Score 779; DB 5; Length 149194;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 4802 AGGCTTCCCTGTAGGACAGAAAAGGCCAAGAGGTAAATAGGCACCTAGTTTCATGAAA 4861
    |||||
DB 86903 AGGCTTCCCTGTAGGACAGAAAAGGCCAAGAGGTAAATAGGCACCTAGTTTCATGAAA 86844
    |||||
QY 4862 TAAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCCAGG 4921
    |||||

```



Db 86843 TAAATCCAGATTCCGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGG 86784  
Qy 4922 CCACAGTAAACCCAGGAGTATCCCAAGGGTTAGGTATACGATATCACTTACACTGGCGCT 4981  
Db 86783 CCACAGTAAACCCAGGAGTATCCCAAGGGTTAGGTATACGATATCACTTACACTGGCGCT 86724  
Qy 4982 GAAGGCCACAGTCCTCAGGGAAGTTCGAGAAATGAATGAAYACTCAAGAGACATCTAA 5041  
Db 86723 GAAGGCCACAGTCCTCAGGGAAGTTCGAGAAATGAATGAAYACTCAAGAGACATCTAA 86664  
Qy 5042 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCCCTGYTCTGTGCTATAGCTTTAAAAA 5101  
Db 86663 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCCCTGYTCTGTGCTATAGCTTTAAAAA 86604  
Qy 5102 GAATCTGCAACTTTCCCAAAAGCAGGACTTTCAGCCCATACGAAATGCTGTATGAAGGC 5161  
Db 86603 GAATCTGCAACTTTCCCAAAAGCAGGACTTTCAGCCCATACGAAATGCTGTATGAAGGC 86544  
Qy 5162 CTTTCATACCAATGACCTTGTGCTTGACCCAGACAGCCAACTTAGTTGAGACATCAC 5221  
Db 86543 CTTTCATACCAATGACCTTGTGCTTGACCCAGACAGCCAACTTAGTTGAGACATCAC 86484  
Qy 5222 CTCCTTAGCCAAATCAACAAGTCTTAAACATTTACAAGGAACTATCCCTGAGAAGA 5281  
Db 86483 CTCCTTAGCCAAATCAACAAGTCTTAAACATTTACAAGGAACTATCCCTGAGAAGA 86424  
Qy 5282 GGGAAAGAACTATTCCACCCWGTGACATGTTATAGTCAAGTCCCTTCCTCTAAATTC 5341  
Db 86423 GGGAAAGAACTATTCCACCCWGTGACATGTTATAGTCAAGTCCCTTCCTCTAAATTC 86364  
Qy 5342 CCATCCCTTAGATACATCTGGGAAGGACCTACCCAGTCAATTTATATACCCCAACTGC 5401  
Db 86363 CCATCCCTTAGATACATCTGGGAAGGACCTACCCAGTCAATTTATATACCCCAACTGC 86304  
Qy 5402 GGTAAAGTGGCTGAGTGGAGTCTTGGATACATCACATCTGAGTCAAACTCTGGATACT 5461  
Db 86303 GGTAAAGTGGCTGAGTGGAGTCTTGGATACATCACATCTGAGTCAAACTCTGGATACT 86244  
Qy 5462 GCCAAAGAACTGAAATCCAGGAGACAAACCTAGCTATTCCTGTGAACTCTTAGAGGA 5521  
Db 86243 GCCAAAGAACTGAAATCCAGGAGACAAACCTAGCTATTCCTGTGAACTCTTAGAGGA 86184  
Qy 5522 TTTGGCCCTGCTCTTCAACAACAACAGGAGGAAAGTAACTAAATCATTAATCCCCA 5581  
Db 86183 TTTGGCCCTGCTCTTCAACAACAACAGGAGGAAAGTAACTAAATCATTAATCCCCA 86125  
Qy 5582 TGGSCCTCCCTTATCATATTTTCTCTKTASTTSTTAACTCTTCTCACTCTCACTG 5641  
Db 86124 TGGSCCTCCCTTATCATATTTTCTCTKTASTTSTTAACTCTTCTCACTCTCACTG 86065  
Qy 5642 CACCCCTCCATGCGCTGTATGACAGTAGTCCCTTACCMAGAGTTTCTATGGAGAA 5701  
Db 86064 CACCCCTCCATGCGCTGTATGACAGTAGTCCCTTACCAAGAGTTTCTATGGAGAA 86005  
Qy 5702 TGCAGGTCGCGAAATATGATGCCCATGATAGGAGTCTTSTTAAGGAAACCCCA 5761  
Db 86004 TGCAGGTCGCGAAATATGATGCCCATGATAGGAGTCTTSTTAAGGAAACCCCA 85945  
Qy 5762 CTTCACTGCCACACCCATATGCCCGCACTGCTATCACTCTGCACTCTTTGCGATGC 5821  
Db 85944 CTTTCACTGCCACACCCATATGCCCGCACTGCTATCACTCTGCACTCTTTGCGATGC 85885  
Qy 5822 ATGCAAAATCTCATTTATGGACAGGAAATGATTAATCCTAGTTGTCTCGAGGACTTG 5881  
Db 85884 ATGCAAAATCTCATTTATGGACAGGAAATGATTAATCCTAGTTGTCTCGAGGACTTG 85825  
Qy 5882 GAGTCACTCTGTGGACTTACTTCAACCAACTGGTATGCTGATGGGGTGGAGTTC 5941  
Db 85824 GAGTCACTCTGTGGACTTACTTCAACCAACTGGTATGCTGATGGGGTGGAGTTC 85765  
Qy 5942 AAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTATCTCCCAACTCACCSSGGTACATG 6001  
Db 85764 AAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTATCTCCCAACTCACCSSGGTACATG 85705

Qy 6002 GCACCTTAGCCCTCAAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCCGTACCC 6061  
Db 85704 GCACCTTAGCCCTCAAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCCGTACCC 85645  
Qy 6062 ATACTCGCCTCGTAAAGCCCTATTTAATACCAACCTCACTGGGCTCCATGAGGTCTCGGCC 6121  
Db 85644 ATACTCGCCTCGTAAAGCCCTATTTAATACCAACCTCACTGGGCTCCATGAGGTCTCGGCC 85585  
Qy 6122 AAAACCCCTACTAACTGTGGATATGCCCTCCCTGAACTTCARGCCATATGTTTCAAATCC 6181  
Db 85584 AAAACCCCTACTAACTGTGGATATGCCCTCCCTGAACTTCARGCCATATGTTTCAAATCC 85525  
Qy 6182 CTGTACCTGGAACATGGAACAACTTCAGACAGAAATAAAACCACTTCCGTTTGTAGTAG 6241  
Db 85524 CTGTACCTGGAACATGGAACAACTTCAGACAGAAATAAAACCACTTCCGTTTGTAGTAG 85465  
Qy 6242 GACCTCTTGTTCCTCAATSTGGAATAACCCATACCTCAAACTCACTGTGTAAAATTTA 6301  
Db 85464 GACCTCTTGTTCCTCAATSTGGAATAACCCATACCTCAAACTCACTGTGTAAAATTTA 85405  
Qy 6302 GCAATACATACACAAACCACTCCCAATGCATCAGGTGGGTAACTCCTCCCAACAAA 6361  
Db 85404 GCAATACATACACAAACCACTCCCAATGCATCAGGTGGGTAACTCCTCCCAACAAA 85345  
Qy 6362 TAGTCTGCCTACCTCAGGAATATTTTGTCTGTGTGCTACCTCAGCCCTATCGTTGTTGA 6421  
Db 85344 TAGTCTGCCTACCTCAGGAATATTTTGTCTGTGTGCTACCTCAGCCCTATCGTTGTTGA 85285  
Qy 6422 ATGGCTCTTCAAGATCTATGTGCTCTCTCATCTTGTAGTCCCTCACTGRCATCTACA 6481  
Db 85284 ATGGCTCTTCAAGATCTATGTGCTCTCTCATCTTGTAGTCCCTCACTGRCATCTACA 85225  
Qy 6482 CTGAACAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCATTTC 6541  
Db 85224 CTGAACAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCATTTC 85165  
Qy 6542 TTCTTTTGTATAGGACAGAGTGTAGTGCACTAGGTACTGGCATTTGGCGGTATCA 6601  
Db 85164 TTCTTTTGTATAGGACAGAGTGTAGTGCACTAGGTACTGGCATTTGGCGGTATCA 85105  
Qy 6602 CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAATGGGACATGGAAACGG 6661  
Db 85104 CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAATGGGACATGGAAACGG 85045  
Qy 6662 TCGCGACTCCCTGTGTCACCTTGGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTC 6721  
Db 85044 TCGCGACTCCCTGTGTCACCTTGGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTC 84985  
  
RESULT 22  
AF156963  
LOCUS  
DEFINITION Homo sapiens human endogenous retrovirus W envC7-1 envelope protein (env) gene, complete cds.  
ACCESSION AF156963  
VERSION AF156963.1 GI:8272467  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
REFERENCE 1 (bases 1 to 1617)  
AUTHORS Voisset,C., Bouton,O., Bedin,F., Duret,L., Mandrand,B., Mallet,F. and Paranhos-Baccala,G.  
TITLE Chromosomal distribution and coding capacity of the human endogenous retrovirus HERV-W family  
JOURNAL AIDS Res. Hum. Retroviruses 16 (8), 731-740 (2000)  
PUBMED 10826480  
REFERENCE 2 (bases 1 to 1617)  
AUTHORS Voisset,C., Bouton,O., Bedin,F., Duret,L., Mandrand,B., Mallet,F. and Paranhos-Baccala,G.

TITLE Direct Submission  
JOURNAL Submitted (09-JUN-1999) UMR103 CNRS-bioMerieux, ENS Lyon, 46 allée  
d'Italie, Lyon 69384, France  
FEATURES  
source  
1. .1617  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="7"  
/clone="envC7-1"  
/notes="human/rodent monochromosomal NIGMS somatic hybrid  
panel n#2 (Coriell Institute, Camden, NJ, USA); clone  
GM/NA10791 retaining human chromosome 7"  
gene  
1. .1617  
/gene="env"  
CDS  
1. .1617  
/gene="env"  
/codon\_start=1  
/product="envelope protein"  
/protein\_id="AA874215.1"  
/db\_xref="GI:8272468"  
/translations="MALPHYIFLFTVLLPSFTLTAPPPCRMTSSPYOEFRLWRMP  
GNIDAFSTSLSGTFTTAHMPNRCVHSALCMHANTHYWTGMINPSCPGGLGV  
TNCWTYFTOTGMDGGVOOAREKHKEVVISOLTRVHGTSYPYKGLDLSKLEHETRT  
HTRLSLENTLTGLHEVSAQNPNCWICLPLNRPVVSIPVPEONNPFSTEINTSV  
LVGLVSNLEITHSNLTCKVFNSTYTNQCIKWPTPTQIVLPSPGFFVCGTSA  
YRCNGSSSMCFSLFLVPPMTIYTQDDLYSYVISKPRNKRVPILPFIAGVIGALG  
TGIGGTTSTFYKLSQELNGDMERVADSLVTLQDLNSLAALVLRNRRALLTLAE  
RGCTCLFEECYQNSQIGIVTEKVEIRDRIOORAEELNRTGPMGLLSOMWPTLP  
FLGLAAIILLLLFPCIFENLLVFNVSRIEAVKLQWEPKMQSKTIYRPLDRPASP  
RSDVNDIKGTPPEISAAQPLLRPSNGSS"  
repeat\_region  
1. .1617  
/rpt\_family="human endogenous retrovirus W"  
primer\_bind  
primer\_bind  
1. .18  
1597. .1617  
ORIGIN  
Query Match 9.6%; Score 731; DB 5; Length 1617;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 5685 AGAGTTTCTATGAGAGATGAGCGTCCCGGAAATATGATGCCCATCGTATAGGAGTCT 5744  
DB 105 AGAGTTTCTATGAGAGATGAGCGTCCCGGAAATATGATGCCCATCGTATAGGAGTCT 164  
QY 5745 TTSTAAGGGAACCCACCTTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTC 5804  
DB 165 TTCTAAGGGAACCCACCTTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTC 224  
QY 5805 TGCACCTCTTTGATGATGATCAATATCTCAATATTGGACAGGAAATGATTAATCTAG 5864  
DB 225 TGCACCTCTTTGATGATGATCAATATCTCAATATTGGACAGGAAATGATTAATCTAG 284  
QY 5865 TTGCTCTGGAGGACTTGGAGTCACTGTCTTTGGACTTACTTACCACAACTCGGTATGC 5924  
DB 285 TTGCTCTGGAGGACTTGGAGTCACTGTCTTTGGACTTACTTACCACAACTCGGTATGC 344  
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAGAGNAGTATCTCCCA 5984  
DB 345 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAGAGNAGTATCTCCCA 404  
QY 5985 ACTCACCGGGTACATGGCACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACA 6044  
DB 405 ACTCACCGGGTACATGGCACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACA 464  
QY 6045 TGAACCCCTCGTACCATACCTCGCTGGTGAAGCTATTATTAACCACTCTCTGGCT 6104  
DB 465 TGAACCCCTCGTACCATACCTCGCTGGTGAAGCTATTATTAACCACTCTCTGGCT 524  
QY 6105 CCATGAGGTTCTGGCCCAAAACCTTACTACTTGTGGATATGCTCCCTGCACTTCAR 6164  
DB 525 CCATGAGGTTCTGGCCCAAAACCTTACTACTTGTGGATATGCTCCCTGCACTTCAR 584

QY 6165 GCATATGTTTCAATCCCTGTACTGAAACAATGGAACAACCTTCAGCACAGAAATAAACAC 6224  
DB 585 GCATATGTTTCAATCCCTGTACTGAAACAATGGAACAACCTTCAGCACAGAAATAAACAC 644  
QY 6225 CACTTCGGTTTATAGTAGGACCTCTCTGTTTCAATGGAATAACCAATGCACTCAACCT 6284  
DB 645 CACTTCGGTTTATAGTAGGACCTCTCTGTTTCAATGGAATAACCAATGCACTCAACCT 704  
QY 6285 CACTCTGTAAAAATTTAGCAATACTACATACACAACAACCACTCCCAATGCACTCAGGTGGT 6344  
DB 705 CACTCTGTAAAAATTTAGCAATACTACATACACAACAACCACTCCCAATGCACTCAGGTGGT 764  
QY 6345 AACTCTCTCCACACAATAATAGTCTGCCTACCCCTCAGGAATATTTTGTCTGTGTACCTC 6404  
DB 765 AACTCTCTCCACACAATAATAGTCTGCCTACCCCTCAGGAATATTTTGTCTGTGTACCTC 824  
QY 6405 AGCCTATCGTTGTTGAATGGCTCTTCAGAACTATATGCTGCTCTCATTTCTAGTGCC 6464  
DB 825 AGCCTATCGTTGTTGAATGGCTCTTCAGAACTATATGCTGCTCTCATTTCTAGTGCC 884  
QY 6465 CCCATATGRCATCTACACTGAACAAGATTATACAGTTATGCTATCTAAGCCCGCAA 6524  
DB 885 CCTATGACCACTACACTGNAACAAGATTATACAGTTATGCTATCTAAGCCCGCAA 944  
QY 6525 CAAAAGTAGTACCATTCTCTTTTGTATAGGACAGGAGTCTAGGTGCACTAGGTAC 6584  
DB 945 CAAAAGTAGTACCATTCTCTTTTGTATAGGACAGGAGTCTAGGTGCACTAGGTAC 1004  
QY 6585 TGCCATTGGGGTATCACAACCTTCACTAGTCTTACTACTACAACTATCTCAAGAACTAAA 6644  
DB 1005 TGCCATTGGGGTATCACAACCTTCACTAGTCTTACTACTACAACTATCTCAAGAACTAAA 1064  
QY 6645 TGGGGACATGGAACGGTCCCGACTCCCTGGTCACTTGCAGAGATCACTTAACCTCCCT 6704  
DB 1065 TGGGGACATGGAACGGTCCCGACTCCCTGGTCACTTGCAGAGATCACTTAACCTCCCT 1124  
QY 6705 AGCAGCAGTAGTCTCTTC 6721  
DB 1125 AGCAGCAGTAGTCTCTTC 1141  
RESULT 23  
BD196239  
LOCUS  
DEFINITION BD196239 1948 bp DNA linear PAT 17-JUL-2003  
Endogenous retroviral sequences, associated with autoimmune  
diseases and/or with pregnancy disorders.  
ACCESSION BD196239  
VERSION BD196239.1 GI:33006009  
KEYWORDS JP 2002512530-A/5.  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 1948)  
AUTHORS Beseme,F., Blond,J.L., Bouton,O., Mandrand,B. and Mallet,F.  
TITLE Endogenous retroviral sequences, associated with autoimmune  
diseases and/or with pregnancy disorders  
JOURNAL Patent: JP 2002512530-A 5 23-APR-2002;  
COMMENT BIO MERIEUX  
OS Unidentified  
PN JP 2002512530-A/5  
PD 23-APR-2002  
PF 06-JUL-1998 JP 1999508244  
PR 07-JUL-1997 FR 97/08815  
PI FREDERIC BESEME, JEAN LUC BLOND, OLIVIER BOUTON, BERNARD  
MANDRAND,  
PI FRANCOIS MALLET  
PC C12N15/48, C07K14/15, C12Q1/68, C07K16/10, G01N33/569 CC  
Strandedness: Single;  
CC Topology: Linear;  
CC Endogenous retroviral sequences, associated with autoimmune  
diseases  
CC and/or with pregnancy disorders

QY	6585	TGGCATTTGGCGGTATCAACACTTACTAGTTCTTACTACAACTATCTCAAGAACTAAA	6644
DB	1723	TGGCATTTGGCGGTATCAACACTTACTAGTTCTTACTACAACTATCTCAAGAACTAAA	1782
QY	6645	TGGGACATGGNAACGGTCCGCGACTCCCTGGTCACTTCGACAGATCAACTTAATCCCT	6704
DB	1783	TGGGACATGGNAACGGTCCGCGACTCCCTGGTCACTTCGACAGATCAACTTAATCCCT	1842
QY	6705	AGCAGCAGTAGTCCTTC	6721
DB	1843	AGCAGCAGTAGTCCTTC	1859
RESULT 24			
BD267481			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
QY	5685	AGAGTTTCTATGAGAAATGACGCGTCCCGAAATATTGATGCCCATCGTATAGGAGTCT	5744
DB	823	AGAGTTTCTATGAGAAATGACGCGTCCCGAAATATTGATGCCCATCGTATAGGAGTCT	882
QY	5745	TTSTAAGGAAACCCACCTTCACTGCCCCACACCCATATGCCCCGCAACTCTATCACTC	5804
DB	883	TTSTAAGGAAACCCACCTTCACTGCCCCACACCCATATGCCCCGCAACTCTATCACTC	942
QY	5805	TGCCACTCTTTCATGATGCAATTAATCTATTATGGACAGAAAAATGATTAATCCCTAG	5864
DB	943	TGCCACTCTTTCATGATGCAATTAATCTATTATGGACAGAAAAATGATTAATCCCTAG	1002
QY	5865	TTGTCTCGGAGACTTGGAGTCACTGTCTGTGGACTTCTTCAACCCAACTGGTATGTC	5924
DB	1723	TTGTCTCGGAGACTTGGAGTCACTGTCTGTGGACTTCTTCAACCCAACTGGTATGTC	1782
QY	6045	TGAACCTCCGACCAATAGTCTGCCCTACCCCTACCAAGACTAGATCTCTCAAACTACA	6104
DB	1183	TGAACCTCCGACCAATAGTCTGCCCTACCCCTACCAAGACTAGATCTCTCAAACTACA	1242
QY	6105	CCATGAGTCTCGGCGCCAAACCTTAACTACTGTTGGATATGCTCCCTCCCTGAACTTCAR	6164
DB	1243	CCATGAGTCTCGGCGCCAAACCTTAACTACTGTTGGATATGCTCCCTCCCTGAACTTCAG	1302
QY	6165	GCCATATGTTCAATCCCTGTACTGAAACAAATGGAACAACTTCAGACAGAAATAACAC	6224
DB	1303	GCCATATGTTCAATCCCTGTACTGAAACAAATGGAACAACTTCAGACAGAAATAACAC	1362
QY	6225	CACCTCCGTTTATGAGGACCTCTGTTCCCAATSTGGAATAACCCATACCTCAAACT	6284
DB	1363	CACCTCCGTTTATGAGGACCTCTGTTCCCAATSTGGAATAACCCATACCTCAAACT	1422
QY	6285	CACCTGTTGAAAATTTAGCAATACTACATACCAACCACTCCCAATGATCAGTGGGT	6344
DB	1423	CACCTGTTGAAAATTTAGCAATACTACATACCAACCACTCCCAATGATCAGTGGGT	1482
QY	6345	AACCTCCACCAAAATAGTCTGCCCTACCCCTCAGGAATATTTTGTGTGTFGACCTC	6404
DB	1483	AACCTCCACCAAAATAGTCTGCCCTACCCCTCAGGAATATTTTGTGTGTFGACCTC	1542
QY	6405	AGCCTATCGTTTGTGATGGCTCTTCAGAACTATGTCCTTCTCATTTCTTAGTGCC	6464
DB	1543	AGCCTATCGTTTGTGATGGCTCTTCAGAACTATGTCCTTCTCATTTCTTAGTGCC	1602
QY	6465	CCCVATGRCATCTACACTGAAACAGATTTATACAGTTATGTCATATCTAAGCCCCGAA	6524
DB	1603	CCCVATGRCATCTACACTGAAACAGATTTATACAGTTATGTCATATCTAAGCCCCGAA	1662
QY	6525	CAAAAGAGTACCCATCTTCTTTGTTATGAGAGCAGAGTGTAGTGCACTAGGTAC	6584
DB	1663	CAAAAGAGTACCCATCTTCTTTGTTATGAGAGCAGAGTGTAGTGCACTAGGTAC	1722

Db 1003 TTGTCCTGGAGGACTTGGAGTCACGTCTCTGTTGGACTTACTTACCCCAAACTGGTATGTC 1062

Qy 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAAGAAATATCTCCCA 5984

Db 1063 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAAGAAATATCTCCCA 1122

Qy 5985 ACTCACCGGGTACATGCGCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAACTACA 6044

Db 1123 ACTCACCGGGTACATGCGCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAACTACA 1182

Qy 6045 TGAACCCCTCGPACCCATCTCGCCTGGTAAAGCTTATTTAAACCAACCTCTACCTGGCT 6104

Db 1183 TGAACCCCTCGPACCCATCTCGCCTGGTAAAGCTTATTTAAACCAACCTCTACCTGGCT 1242

Qy 6105 CATAGAGTCTCGGCCCCAAAACCTACTACTGTGGATATGCTCCCTGAACTTCAR 6164

Db 1243 CCATAGAGTCTCGGCCCCAAAACCTACTACTGTGGATATGCTCCCTGAACTTCAG 1302

Qy 6165 GCCATATGTTTCAATCCCTGACTGAACAATGGAACAACCTTCAGCACAGAAATAACAC 6224

Db 1303 GCCATATGTTTCAATCCCTGACTGAACAATGGAACAACCTTCAGCACAGAAATAACAC 1362

Qy 6225 CACTCCGTTTTAGTAGGACCTCTGTTTCCAAATSTGGAATTAACCCATACCTCAAACT 6284

Db 1363 CACTCCGTTTTAGTAGGACCTCTGTTTCCAAATSTGGAATTAACCCATACCTCAAACT 1422

Qy 6285 CACTGTGTAATTTAGCAATACATACACAACAACCTCCCAATGCAATCAGGTGGT 6344

Db 1423 CACTGTGTAATTTAGCAATACATACACAACAACCTCCCAATGCAATCAGGTGGT 1482

Qy 6345 AACTCTCCACACAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGGTACCTC 6404

Db 1483 AACTCTCCACACAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGGTACCTC 1542

Qy 6405 AGCTATGTTGTTGAATGGCTCTTCAGAACTATGTGCTCTCTCATTTCTAGTGCC 6464

Db 1543 AGCTATGTTGTTGAATGGCTCTTCAGAACTATGTGCTCTCTCATTTCTAGTGCC 1602

Qy 6465 CCYATGRCATCTACATGAACAAGATTTATACAGTTATGATATCTAAGCCCGCAA 6524

Db 1603 CCYATGRCATCTACATGAACAAGATTTATACAGTTATGATATCTAAGCCCGCAA 1662

Qy 6525 CAAAGAGTACCANTCTCTTTTGTATAGGACAGAGTGTAGGTGCATAGGTAC 6584

Db 1663 CAAAGAGTACCANTCTCTCTTTTGTATAGGACAGAGTGTAGGTGCATAGGTAC 1722

Qy 6585 TGSCATTTGGCGGTATCAACCTCTACTAGTTCTACTACAACTATCTCAAGAACTAAA 6644

Db 1723 TGSCATTTGGCGGTATCAACCTCTACTAGTTCTACTACAACTATCTCAAGAACTAAA 1782

Qy 6645 TGGGACATGGAACGGGTGCGCGACTCCCTGGTCACTTGGCAAGATCAACTTAACCTCCCT 6704

Db 1783 TGGGACATGGAACGGGTGCGCGACTCCCTGGTCACTTGGCAAGATCAACTTAACCTCCCT 1842

Qy 6705 AGCAGCAGTAGTCTCTTC 6721

Db 1843 AGCAGCAGTAGTCTCTTC 1859

RESULT 25

AX000960

LOCUS

DEFINITION

Sequence 5 from Patent WO9902696.

ACCESSION

AX000960

VERSION

AX000960.1 GI:7241202

KEYWORDS

unidentified

SOURCE

unclassified sequences.

ORGANISM

1 (bases 1 to 1948)

REFERENCE

1 (bases 1 to 1948)

AUTHORS

Beseme, F. and Blond, J.

TITLE

ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISEASES OR WITH PREGNANCY DISORDERS

JOURNAL

Patent: WO 9902696-A 5 21-JAN-1999;

FEATURES

Location: BESEME FREDERIC (FR)

source

1. .1948

/organism="unidentified"

/mol\_type="unassigned DNA"

/db\_xref="taxon:32644"

ORIGIN

Query Match 9.6%; Score 731; DB 2; Length 1948;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 5685 AGAGTTTCTATGAGAAATGCGCTCCCGAAATATGATGCCCCATCGTATAGGAGTCT 5744

Db 823 AGAGTTTCTATGAGAAATGCGCTCCCGAAATATGATGCCCCATCGTATAGGAGTCT 882

Qy 5745 TTSTAAGGGAACCCCACTTCACTGCCCCACACCCATATGCCCCCACTGCTATCACC 5804

Db 883 TTSTAAGGGAACCCCACTTCACTGCCCCACACCCATATGCCCCCACTGCTATCACC 942

Qy 5805 TGCCACTCTTTGATGCAATCAATCAATTTGGACAGGAAAAATGATTAATCCTAG 5864

Db 943 TGCCACTCTTTGATGCAATCAATCAATTTGGACAGGAAAAATGATTAATCCTAG 1002

Qy 5865 TTGCTCTGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTACCCAACTGGTATGTC 5924

Db 1003 TTGCTCTGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTACCCAACTGGTATGTC 1062

Qy 5925 TGATGGGGTGGAGTTCAGATCAGGCAAGAGAAAAACATGTAAAAAGAAATATCTCCCA 5984

Db 1063 TGATGGGGTGGAGTTCAGATCAGGCAAGAGAAAAACATGTAAAAAGAAATATCTCCCA 1122

Qy 5985 ACTCACCGGGTACATGCGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACA 6044

Db 1123 ACTCACCGGGTACATGCGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACA 1182

Qy 6045 TGAACCCCTCGTACCCATCTCGCTGGTAAAGCTTATTTAATACCAACCTCTACCTGGCT 6104

Db 1183 TGAACCCCTCGTACCCATCTCGCTGGTAAAGCTTATTTAATACCAACCTCTACCTGGCT 1242

Qy 6105 CATAGAGTCTCGGCCCCAAAACCTACTACTGTTGGATATGCTCCCTGAACTTCAR 6164

Db 1243 CATAGAGTCTCGGCCCCAAAACCTACTACTGTTGGATATGCTCCCTGAACTTCAG 1302

Qy 6165 GCATATGTTTCAATCCCTGACTGAACAATGGAACAACCTTCAGCACAGAAATAACAC 6224

Db 1303 GCATATGTTTCAATCCCTGACTGAACAATGGAACAACCTTCAGCACAGAAATAACAC 1362

Qy 6225 CACTTCCTCCACACAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGGTACCTC 6284

Db 1363 CACTTCCTCCACACAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGGTACCTC 1422

Qy 6285 CACTGTGTAATTTAGCAATACATACACAACAACCTCCCAATGCAATCAGGTGGT 6344

Db 1423 CACTGTGTAATTTAGCAATACATACACAACAACCTCCCAATGCAATCAGGTGGT 1482

Qy 6345 AACTCTCCACACAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGGTACCTC 6404

Db 1483 AACTCTCCACACAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGGTACCTC 1542

Qy 6405 AGCTATGTTGTTGAATGGCTCTTCAGAACTATGTGCTCTCTCATTTCTAGTGCC 6464

Db 1543 AGCTATGTTGTTGAATGGCTCTTCAGAACTATGTGCTCTCTCATTTCTAGTGCC 1602

Qy 6465 CCYATGRCATCTACATGAACAAGATTTATACAGTTATGATATCTAAGCCCGCAA 6524

Db 1603 CCYATGRCATCTACATGAACAAGATTTATACAGTTATGATATCTAAGCCCGCAA 1662

Qy 6525 CAAAGAGTACCANTCTCTTTTGTATAGGACAGAGTGTAGGTGCATAGGTAC 6584

Db 1663 CAAAGAGTACCANTCTCTCTTTTGTATAGGACAGAGTGTAGGTGCATAGGTAC 1722

```
QY 6585 TGGCATTTGGCGGTATCACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 6644
Db 1723 TGGCATTTGGCGGTATCACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 1782
QY 6645 TGGGACATGGAACGGGTGCGGACTCCCTGGTGCACCTTGGCAAGATCAACTTAACTCCCT 6704
Db 1783 TGGGACATGGAACGGGTGCGGACTCCCTGGTGCACCTTGGCAAGATCAACTTAACTCCCT 1842
QY 6705 AGCAGCATGAGTCCTTC 6721
Db 1843 AGCAGCATGAGTCCTTC 1859

RESULT 26
AX027474
LOCUS AX027474 1948 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 24 from Patent FR2788784.
ACCESSION AX027474
VERSION AX027474.1 GI:10188438
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1. Mallet, F., Voisset, C. and Paranhos, B.G.
AUTHORS Mallet, F.; Voisset, C.; Paranhos, B.G.
JOURNAL Patent: FR 2788784-A 24 28-JUL-2000;
BIO MERIEUX (FR)
FEATURES
source
1. 1948
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 9.6%; Score 731; DB 2; Length 1948;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 5685 AGAGTTTCTATGGAGATGACGGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 5744
Db 823 AGAGTTTCTATGGAGATGACGGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 882
QY 5745 TTSTAAGGAACCCCACTTCACTGCCCCACACCCATATGCCCGGCAACTGCTATCACTC 5804
Db 883 TTSTAAGGAACCCCACTTCACTGCCCCACACCCATATGCCCGGCAACTGCTATCACTC 942
QY 5805 TGCCACTCTTTGCATGCAATGCAATCACTATTATTGGACAGGAAATGATTAATCCTAG 5864
Db 943 TGCCACTCTTTGCATGCAATGCAATCACTATTATTGGACAGGAAATGATTAATCCTAG 1002
QY 5865 TTGTCTCTGGAGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCAAACTGGTATGTC 5924
Db 1003 TTGTCTCTGGAGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCAAACTGGTATGTC 1062
QY 5925 TGATGGGGTGGAGTTCAGATCAGCAAGAGAGAAAACATGTAAGAAGTAACTCCCA 5984
Db 1063 TGATGGGGTGGAGTTCAGATCAGCAAGAGAGAAAACATGTAAGAAGTAACTCCCA 1122
QY 5985 ACTCACCGGTACATGGCACTCTAGCCCTACAGGACTAGATCTCTCAAACTACA 6044
Db 1123 ACTCACCGGTACATGGCACTCTAGCCCTACAGGACTAGATCTCTCAAACTACA 1182
QY 6045 TGAACCCCTCCGTACCCATCTCGCTGGTGAAGCCCTATTATTAACACCCCTCACTGGGCT 6104
Db 1183 TGAACCCCTCCGTACCCATCTCGCTGGTGAAGCCCTATTATTAACACCCCTCACTGGGCT 1242
QY 6105 CCATGAGGTCTGGGCCAAAACCCCTACTAACTGTTGGATATGCTCCGCCCTGAATTCAR 6164
Db 1243 CCATGAGGTCTGGGCCAAAACCCCTACTAACTGTTGGATATGCTCCGCCCTGAATTCAG 1302
QY 6165 GCCATATGTTTCAATCCCTGTACCTGAACTGAAAGAACTTTCAGCAGAGAAATAACAC 6224
```

```
Db 1303 GCCATATGTTTCAATCCCTGTACCTGAACTGAACTCTCAGCAGAGAAATAACAC 1362
QY 6225 CACTTCCGTTTTAGTAGGACCTCTTGTTCCTCAATSTGGAATAAACCCATACCTCAAACT 6284
Db 1363 CACTTCCGTTTTAGTAGGACCTCTTGTTCCTCAATSTGGAATAAACCCATACCTCAAACT 1422
QY 6285 CACTGTGTAAATTTAGCAATATACATACACAACTCCCAATGATCAGGTGGT 6344
Db 1423 CACTGTGTAAATTTAGCAATATACATACACAACTCCCAATGATCAGGTGGT 1482
QY 6345 AACTCTCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTTGTCTGTGTACCTC 6404
Db 1483 AACTCTCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTTGTCTGTGTACCTC 1542
QY 6405 AGCTATCGTTTGTGAATGGCTCTTCAAGATCTAATGTGCTCTCTCAATTTCTAGTGC 6464
Db 1543 AGCTATCGTTTGTGAATGGCTCTTCAAGATCTAATGTGCTCTCTCAATTTCTAGTGC 1602
QY 6465 CCCATGRCCTACACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAA 6524
Db 1603 CCCATGRCCTACACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAA 1662
QY 6525 CAAAGAGTACCCATCTTCTCTTTTATAGGAGCAGGAGTCTAGTGCCTAGGAGTAC 6584
Db 1663 CAAAGAGTACCCATCTTCTCTTTTATAGGAGCAGGAGTCTAGTGCCTAGGAGTAC 1722
QY 6585 TGGCATTTGGCGGTATCACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 6644
Db 1723 TGGCATTTGGCGGTATCACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 1782
QY 6645 TGGGACATGGAACGGGTGCGGACTCCCTGGTGCACCTTGGCAAGATCAACTTAACTCCCT 6704
Db 1783 TGGGACATGGAACGGGTGCGGACTCCCTGGTGCACCTTGGCAAGATCAACTTAACTCCCT 1842
QY 6705 AGCAGCATGAGTCCTTC 6721
Db 1843 AGCAGCATGAGTCCTTC 1859

RESULT 27
AX027505
LOCUS AX027505 1948 bp mRNA linear PRI 10-FEB-1999
DEFINITION Homo sapiens endogenous retrovirus W envelope protein mRNA, partial cds.
ACCESSION AF072505
VERSION AF072505.1 GI:4262289
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1. (bases 1 to 1948)
AUTHORS Blond, J.L., Beseme, F., Duret, L., Bouton, O., Bedin, F., Perron, H., Mandrand, B., and Mallet, F.
TITLE Molecular characterization and placental expression of HERV-W, a new human endogenous retrovirus family
JOURNAL J. Virol. 73 (2), 1175-1185 (1999)
PUBMED 982319
REFERENCE 2. (bases 1 to 1948)
AUTHORS Blond, J.-L., Beseme, F. and Mallet, F.
TITLE Direct Submission
JOURNAL Submitted (19-JUN-1998) UMI03 CNRS-bioMerieux, ENS Lyon, 46 allée d'Italie, Lyon, Cedex 07 69364, France
FEATURES
source
1. 1948
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="cl.24.4"
/tissue_type="placenta"
/clone_lib="Clontech 5'-stretch plus library, Cat number
```



HL5014a"  
1..1948  
/rpt\_family="HERV-W"  
5'UTR <1..6  
/note="R"  
7..460  
/note="U5"  
464..481  
/note="putative primer binding site for tRNA-W"  
misc\_feature 491..492  
/note="putative splice junction"  
719..>1948  
/codon\_start=1  
/product="envelope protein"  
/protein\_id="AADI454.1"  
/db\_xref="GI:4262290"  
/translation="MALPYHIFLVTVSPSFTLTAPPCKRMTSSSPYQBFELWRMRP  
GNIDATPSYSLSGKTPTFTATHMPNCHVSHATLCMHANTHYWTGKMINPSCPGGLGV  
TCWYTFQTDGGGVQDQAREKHVEIISQLTRVHGTSYKGLDLSKLHETLRT  
HRLVLENTLLTGLHVSQNPNCWICLPLNRPVSI PVPEQNNFSTEINTSV  
LVGPLNSLELTHSLNLCVKFSNTYTTNSQIRWPTPTQIVCLPSGIFVCGHSA  
YRCLNASSSMCFSLFVLPMAIYTBODLYSYVISKPRNKRVPILPFGVGLGALG  
TGGIGTTSTQFYKLSQELNGDMERVADSLVTLQDLNSLAALAVLQNRALDLTAE  
RGGTCFLGEECCYYNQ"

## ORIGIN

Query Match 9.6%; Score 731; DB 5; Length 1948;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGAGAAATGAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 5744  
DB |||||||  
QY 823 AGAGTTTCTATGAGAAATGAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 882  
DB |||||||  
QY 5745 TTSTAAGGGAAACCCCACTTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTC 5804  
DB |||||||  
QY 883 TTGTAAGGAAACCCCACTTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTC 942  
DB |||||||  
QY 5805 TGCACCTTTTGATGATGATCAATATCTCAATATGGACAGGAAAATGATTAACTCTAG 5864  
DB |||||||  
QY 943 TGGCACTTTTGATGATGATCAATATCTCAATATGGACAGGAAAATGATTAACTCTAG 1002  
DB |||||||  
QY 5865 TTGCTCTGGAGGACTTGGAGTCACTGCTGTTGGACTTACTTACCCAACTGGTATGTC 5924  
DB |||||||  
QY 1003 TTGCTCTGGAGGACTTGGAGTCACTGCTGTTGGACTTACTTACCCAACTGGTATGTC 1062  
DB |||||||  
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAACATGTAAAAGAACTATCTCCCA 5984  
DB |||||||  
QY 1063 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAACATGTAAAAGAACTATCTCCCA 1122  
DB |||||||  
QY 5985 ACTCACCGGGTACATGGCACTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACA 6044  
DB |||||||  
QY 1123 ACTCACCGGGTACATGGCACTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACA 1182  
DB |||||||  
QY 6045 TGAACCCCTCGTACCCATCTGCTGTTAGCCCTATTATACACCCCTCTAGTGGCT 6104  
DB |||||||  
QY 1183 TGAACCCCTCGTACCCATCTGCTGTTAGCCCTATTATTAATACCACTCTAGTGGCT 1242  
DB |||||||  
QY 6105 CCATGAGGTTCTGGGCCCAAAACCTTACTACTGTTGATATGCTCCCTGAACTTCAR 6164  
DB |||||||  
QY 1243 CCATGAGGTTCTGGGCCCAAAACCTTACTACTGTTGATATGCTCCCTGAACTTCAG 1302  
DB |||||||  
QY 6165 GCCATATGTTTCAATCCCTGTACTGAACAATGGAAACAACTTCAGACAGAAATAAACAC 6224  
DB |||||||  
QY 1303 GCCATATGTTTCAATCCCTGTACTGACAAATGGAAACAACTTCAGACAGAAATAAACAC 1362  
DB |||||||  
QY 6225 CACTTCGGTTTATGAGGACTCTTGTTCATSTGGAATAATACCCATACCTCAAACT 6284  
DB |||||||  
QY 1363 CACTTCGGTTTATGAGGACTCTTGTTCATSTGGAATAATACCCATACCTCAAACT 1422  
DB |||||||  
QY 6285 CACTTCGTAATAATTTAGCAATATCATACACAACTCCCAATGATCAGTGGGT 6344  
DB |||||||  
QY 1423 CACTTCGTAATAATTTAGCAATATCATACACAACTCCCAATGATCAGTGGGT 1482  
DB |||||||

QY 6345 AACTCTCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGCTACCTC 6404  
DB |||||||  
QY 1483 AACTCTCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGCTACCTC 1542  
DB |||||||  
QY 6405 AGCCTATCGTGTGTTGAATGGCTCTTCAAGATCTATGTGCTCTCTCATCTTAGTGCC 6464  
DB |||||||  
QY 1543 AGCCTATCGTGTGTTGAATGGCTCTTCAAGATCTATGTGCTCTCTCATCTTAGTGCC 1602  
DB |||||||  
QY 6465 CCCATGRCATCTACACTGAACAAGATTATACAGTTATGTCTATCTAAGCCCGCAA 6524  
DB |||||||  
QY 1603 CCCTATGGCCATCTACACTGAACAAGATTATACAGTTATGTCTATCTAAGCCCGCAA 1662  
DB |||||||  
QY 6525 CAAAAGATACCCATCTTCTCTTTTGTATAGAGAGAGAGTCTAGTGCACATAGGTAC 6584  
DB |||||||  
QY 1663 CAAAAGATACCCATCTTCTCTTTTGTATAGAGAGAGAGTCTAGTGCACATAGGTAC 1722  
DB |||||||  
QY 6585 TGCATTGGCGGTATCACAACCTCTACTACTGTTCTACTCAAACTATCTCAAGAACTAAA 6644  
DB |||||||  
QY 1723 TGCATTGGCGGTATCACAACCTCTACTACTGTTCTACTCAAACTATCTCAAGAACTAAA 1782  
DB |||||||  
QY 6645 TGGGGACATGGAACGGGTGCGGACTCCCTGGTCACTTGCAGATCAACTTAACCTCCCT 6704  
DB |||||||  
QY 1783 TGGGGACATGGAACGGGTGCGGACTCCCTGGTCACTTGCAGATCAACTTAACCTCCCT 1842  
DB |||||||  
QY 6705 AGCAGCAGTAGTCTTCTTC 6721  
DB |||||||  
QY 1843 AGCAGCAGTAGTCTTCTTC 1859  
DB |||||||

## RESULT 28

BD221827  
LOCUS  
DEFINITION  
Nucleic sequence and deduced protein sequence family with human  
endogenous retroviral motifs, and their uses.  
ACCESSION  
BD221827  
VERSION  
BD221827.1 GI:33031597  
KEYWORDS  
JP 2002518051-A/22.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 2055)  
AUTHORS  
Alliel,P.M., Perin,J.P. and Rieger,F.  
TITLE  
Nucleic sequence and deduced protein sequence family with human  
endogenous retroviral motifs, and their uses  
JOURNAL  
Patent: JP 2002518051-A 22 25-JUN-2002;  
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE  
COMMENT  
OS Homo sapiens (human)  
PN JP 2002518051-A/22  
PD 25-JUN-2002  
PP 23-JUN-1999 JP 2000556036  
PR 23-JUN-1998 FR 98/07920  
PI PATRICK M ALLIEL,JEAN PIERRE PERIN,FRANCOIS RIEGER PC  
C12N15/09,A01K67/027,A61K31/711,A61K39/21,A61K48/00,A61P21/00,PC  
A61P25/00,  
PC A61P37/06,C07K14/15,C12Q1/68,C12Q1/70,C12N15/00 CC Nucleic  
sequence and deduced protein sequence family with CC  
human endogenous  
CC retroviral motifs, and their uses  
FH Key Location/Qualifiers  
FT CDS (1)..(2055).  
FEATURES  
Location/Qualifiers  
source  
1..2055  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 9.6%; Score 731; DB 2; Length 2055;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
QY 5685 AGAGTTTCTATGGAGAAATGCGCGTCCGGAATATTGATGCCCATCGTATAGGAGTCT 5744
Db 495 AGAGTTTCTATGGAGAAATGCGCGTCCGGAATATTGATGCCCATCGTATAGGAGTCT 554
QY 5745 TTSTAAGGGAACCCCACTTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTC 5804
Db 555 TTCTAAGGGAACCCCACTTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTC 614
QY 5805 TGCCACTCTTTGCATGCAATGCAATCACTATTATTGGACAGAAAAATGATTAATCCTAG 5864
Db 615 TGCCACTCTTTGCATGCAATGCAATCACTATTATTGGACAGAAAAATGATTAATCCTAG 674
QY 5865 TTGTCTCTGAGGACCTTGAGTCACTGCTGTTGGACTTACTTCAACCCAACTGGTATGTC 5924
Db 675 TTGTCTCTGAGGACCTTGAGTCACTGCTGTTGGACTTACTTCAACCCAACTGGTATGTC 734
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAAAGTAATCTCCCA 5984
Db 735 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAAAGTAATCTCCCA 794
QY 5985 ACTCACCGGGTACATGGCACTCTAGCCCCCTACAAAGACTAGATCTCTCAAAACTACA 6044
Db 795 ACTCACCGGGTACATGGCACTCTAGCCCCCTACAAAGACTAGATCTCTCAAAACTACA 854
QY 6045 TGAAACCTCCGTACGCCATCTCGCTGGTGTAGCCCTATTATTATACCACTCCTGGGCT 6104
Db 855 TGAAACCTCCGTACGCCATCTCGCTGGTGTAGCCCTATTATTATACCACTCCTGGGCT 914
QY 6105 CCATGAGGTCTCGGCCAAACCCCTACTAACTGTTGGATATGCTCCCTCGAACTTCAR 6164
Db 915 CCATGAGGTCTCGGCCAAACCCCTACTAACTGTTGGATATGCTCCCTCGAACTTCAG 974
QY 6165 GCCATATGTTTCAATCCCTGTACTGAAACAATGGAACAACCTTCAGCAGAAAAATAAAC 6224
Db 975 GCCATATGTTTCAATCCCTGTACTGAAACAATGGAACAACCTTCAGCAGAAAAATAAAC 1034
QY 6225 CACTTCCGTTTGTAGTAGGACCTTGTTCGAATSTGGAAATAACCCATACCTCAAACT 6284
Db 1035 CACTTCCGTTTGTAGTAGGACCTTGTTCGAATSTGGAAATAACCCATACCTCAAACT 1094
QY 6285 CACTGTGTAAAAATTTAGCAATACATACACAACTCCCACTCCCAATGATCAGGTGGGT 6344
Db 1095 CACTGTGTAAAAATTTAGCAATACATACACAACTCCCACTCCCAATGATCAGGTGGGT 1154
QY 6345 AACTCTCTCCACACAAAATGCTGCTACCTCAGCAATATTTTGTGTGTTGTTGTTGTTG 6404
Db 1155 AACTCTCTCCACACAAAATGCTGCTACCTCAGCAATATTTTGTGTGTTGTTGTTGTTG 1214
QY 6405 AGCCTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 6464
Db 1215 AGCCTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1274
QY 6465 CCATATGRCATCTACACTGAAACAAGATTATACAGTTATGTCATATCTAAGCCCCGCAA 6524
Db 1275 CCATATGRCATCTACACTGAAACAAGATTATACAGTTATGTCATATCTAAGCCCCGCAA 1334
QY 6525 CAAAAGAGTACCAATCTTCTCTTTTGTATTAGGAGCAGGAGTGTAGGTGCACTAGGTAC 6584
Db 1335 CAAAAGAGTACCAATCTTCTCTTTTGTATTAGGAGCAGGAGTGTAGGTGCACTAGGTAC 1394
QY 6585 TGGCAATGCGGTATACAACTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAA 6644
Db 1395 TGGCAATGCGGTATACAACTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAA 1454
QY 6645 TGGGACATGGAACGGGTGCGGACTCCCTGTTGTCACCTTGGCAAGATCAACTTAACTCCCT 6704
Db 1455 TGGGACATGGAACGGGTGCGGACTCCCTGTTGTCACCTTGGCAAGATCAACTTAACTCCCT 1514
QY 6705 AGCAGCAGTGTCTTTC 6721
Db 1515 AGCAGCAGTGTCTTTC 1531
```

```
RESULT 29
AR699776 AR699776 2055 bp DNA linear PAT 14-SEP-2005
LOCUS Sequence 22 from patent US 6919438.
DEFINITION AR699776
ACCESSION AR699776
VERSION AR699776.1 GI:75205567
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 2055)
AUTHORS Alliel,P.M., Perin,J.-P. and Rieger,P.
TITLE Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses
JOURNAL Patent: US 6919438-A 22 19-JUL-2005;
Institut National de la Sante et de la Recherche Medical-Inserm;
Paris;
FRX;
FEATURES Location/Qualifiers
source 1..2055
/mol_type="genomic DNA"
ORIGIN
Query Match 9.6%; Score 731; DB 2; Length 2055;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 5685 AGAGTTTCTATGGAGAAATGCGCGTCCGGAATATTGATGCCCATCGTATAGGAGTCT 5744
Db 495 AGAGTTTCTATGGAGAAATGCGCGTCCGGAATATTGATGCCCATCGTATAGGAGTCT 554
QY 5745 TTSTAAGGGAACCCCACTTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTC 5804
Db 555 TTCTAAGGGAACCCCACTTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTC 614
QY 5805 TGCCACTCTTTGCATGCAATGCAATCACTATTATTGGACAGAAAAATGATTAATCCTAG 5864
Db 615 TGCCACTCTTTGCATGCAATGCAATCACTATTATTGGACAGAAAAATGATTAATCCTAG 674
QY 5865 TTGTCTCTGAGGACCTTGAGTCACTGCTGTTGGACTTACTTCAACCCAACTGGTATGTC 5924
Db 675 TTGTCTCTGAGGACCTTGAGTCACTGCTGTTGGACTTACTTCAACCCAACTGGTATGTC 734
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAAAGTAATCTCCCA 5984
Db 735 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAAAGTAATCTCCCA 794
QY 5985 ACTCACCGGGTACATGGCACTCTAGCCCCCTACAAAGACTAGATCTCTCAAAACTACA 6044
Db 795 ACTCACCGGGTACATGGCACTCTAGCCCCCTACAAAGACTAGATCTCTCAAAACTACA 854
QY 6045 TGAACCCCTCCGTACGCCATCTCGCTGGTGTAGCCCTATTATTATACCACTCCTGGGCT 6104
Db 855 TGAACCCCTCCGTACGCCATCTCGCTGGTGTAGCCCTATTATTATACCACTCCTGGGCT 914
QY 6105 CCATGAGGTCTCGGCCAAACCCCTACTAACTGTTGGATATGCTCCCTCGAACTTCAR 6164
Db 915 CCATGAGGTCTCGGCCAAACCCCTACTAACTGTTGGATATGCTCCCTCGAACTTCAG 974
QY 6165 GCCATATGTTTCAATCCCTGTACTGAAACAATGGAACAACCTTCAGCAGAAAAATAAAC 6224
Db 975 GCCATATGTTTCAATCCCTGTACTGAAACAATGGAACAACCTTCAGCAGAAAAATAAAC 1034
QY 6225 CACTTCCGTTTGTAGTAGGACCTTGTTCGAATSTGGAAATAACCCATACCTCAAACT 6284
Db 1035 CACTTCCGTTTGTAGTAGGACCTTGTTCGAATSTGGAAATAACCCATACCTCAAACT 1094
QY 6285 CACTGTGTAAAAATTTAGCAATACATACACAACTCCCACTCCCAATGATCAGGTGGGT 6344
Db 1095 CACTGTGTAAAAATTTAGCAATACATACACAACTCCCACTCCCAATGATCAGGTGGGT 1154
QY 6345 AACTCTCTCCACACAAAATGCTGCTACCTCAGCAATATTTTGTGTGTTGTTGTTGTTG 6404
Db 1155 AACTCTCTCCACACAAAATGCTGCTACCTCAGCAATATTTTGTGTGTTGTTGTTGTTG 1214
QY 6405 AGCCTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 6464
Db 1215 AGCCTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1274
QY 6465 CCATATGRCATCTACACTGAAACAAGATTATACAGTTATGTCATATCTAAGCCCCGCAA 6524
Db 1275 CCATATGRCATCTACACTGAAACAAGATTATACAGTTATGTCATATCTAAGCCCCGCAA 1334
QY 6525 CAAAAGAGTACCAATCTTCTCTTTTGTATTAGGAGCAGGAGTGTAGGTGCACTAGGTAC 6584
Db 1335 CAAAAGAGTACCAATCTTCTCTTTTGTATTAGGAGCAGGAGTGTAGGTGCACTAGGTAC 1394
QY 6585 TGGCAATGCGGTATACAACTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAA 6644
Db 1395 TGGCAATGCGGTATACAACTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAA 1454
QY 6645 TGGGACATGGAACGGGTGCGGACTCCCTGTTGTCACCTTGGCAAGATCAACTTAACTCCCT 6704
Db 1455 TGGGACATGGAACGGGTGCGGACTCCCTGTTGTCACCTTGGCAAGATCAACTTAACTCCCT 1514
QY 6705 AGCAGCAGTGTCTTTC 6721
Db 1515 AGCAGCAGTGTCTTTC 1531
```

```
QY 6345 AACCTCTCCACACAAATAGTCTGCCCTACCTCAGGAATATTTTGTCTGTGTACTCTC 6404
Db 1155 AACCTCTCCACACAAATAGTCTGCCCTACCTCAGGAATATTTTGTCTGTGTACTCTC 1214
QY 6405 AGCTATCGTGTGTTGAATGGCTCTTCAGAACTCTATGTGCTTCTCTCATTTCTTAGTGCC 6464
Db 1215 AGCTATCGTGTGTTGAATGGCTCTTCAGAACTCTATGTGCTTCTCTCATTTCTTAGTGCC 1274
QY 6465 CCCATGRCATCTACACTGAACAGATTTATACAGTTATGTCTATCTAAGCCCGCAA 6524
Db 1275 CCCTATGACCATCTACACTGAACAGATTTATACAGTTATGTCTATCTAAGCCCGCAA 1334
QY 6525 CAAAGAGTAGCCCATCTTCTCTTTGTTATAGGAGCAGGAGTGCTAGCTAGCTAGGTAC 6584
Db 1335 CAAAGAGTAGCCCATCTTCTCTTTGTTATAGGAGCAGGAGTGCTAGCTAGCTAGGTAC 1394
QY 6585 TGGCATTGGCGGTATCAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 6644
Db 1395 TGGCATTGGCGGTATCAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 1454
QY 6645 TGGGACATGGAACGGGTCCGCACTCCCTGCTGTCACCTTGCAGAGTCAACTTAACCTCT 6704
Db 1455 TGGGACATGGAACGGGTCCGCACTCCCTGCTGTCACCTTGCAGAGTCAACTTAACCTCT 1514
QY 6705 AGCAGCAGTAGTCCTTC 6721
Db 1515 AGCAGCAGTAGTCCTTC 1531

RESULT 30
AX007999
LOCUS AX007999 2055 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 22 from Patent WO967395.
ACCESSION AX007999
VERSION AX007999.1 GI:9995696
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE
1 Perin,J.P., Rieger,F. and Alliel,P.M.
AUTHORS Nucleic sequence and deduced protein sequence family with human
TITLE endogenous retroviral motifs, and their uses
JOURNAL Patent: WO 9967395-A 22 29-DEC-1999;
INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER
FRANCOIS (FR); ALLIEL PATRICK M (FR)
FEATURES
source
1..2055
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
<1..87
/notes="unnamed protein product"
/codon_start=1
/translation="PKTNLVADITSLAKVQVILKTLQCTYP"
CDS
Query Match 9.6%; Score 731; DB 2; Length 2055;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
ORIGIN
QY 5685 AGAGTTTCTATGAGAAATGAGCGTCCGGAATATTTATGATGCCCATCGTATAGGAGTCT 5744
Db 495 AGAGTTTCTATGAGAAATGAGCGTCCGGAATATTTATGATGCCCATCGTATAGGAGTCT 554
QY 5745 TTSTAAGGGAACCCCACTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 5804
Db 555 TTCTAAGGGAACCCCACTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 614
```

Search completed: August 2, 2006, 22:56:40  
Job time : 27465 secs

**This Page Blank (uspto)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2006, 09:10:41 ; Search time 1251 Seconds

(without alignments)  
11340.325 Million cell updates/sec

Title: US-10-717-580-11

Perfect score: 7582

Sequence: 1 caacaatcggtataaac.....tattaaatctgcacttgc 7582

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 1403666 seqs, 935554401 residues

Word size : 6

Total number of hits satisfying chosen parameters: 2781113

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 90 summaries

Database : Issued Patents NA:\*

- 1: /EMC Celerra\_SIDS3/ptodata/2/ina/1\_COMB.seq:\*
- 2: /EMC Celerra\_SIDS3/ptodata/2/ina/5\_COMB.seq:\*
- 3: /EMC Celerra\_SIDS3/ptodata/2/ina/6A\_COMB.seq:\*
- 4: /EMC Celerra\_SIDS3/ptodata/2/ina/6B\_COMB.seq:\*
- 5: /EMC Celerra\_SIDS3/ptodata/2/ina/7\_COMB.seq:\*
- 6: /EMC Celerra\_SIDS3/ptodata/2/ina/H\_COMB.seq:\*
- 7: /EMC Celerra\_SIDS3/ptodata/2/ina/PCUTUS\_COMB.seq:\*
- 8: /EMC Celerra\_SIDS3/ptodata/2/ina/PP\_COMB.seq:\*
- 9: /EMC Celerra\_SIDS3/ptodata/2/ina/RE\_COMB.seq:\*
- 10: /EMC Celerra\_SIDS3/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	779	10.3	2599	US-09-719-554-1	Sequence 1, Appli
2	779	10.3	8523	US-09-573-080A-21	Sequence 21, Appl
3	779	10.3	8523	US-09-854-867-21	Sequence 21, Appl
4	779	10.3	10499	US-09-719-554-3	Sequence 3, Appli
5	779	10.3	13537	US-09-949-016-17382	Sequence 22, Appl
6	731	9.6	2055	US-09-719-554-22	Sequence 5640, Ap
7	731	9.6	2763	US-09-949-016-5640	Sequence 3, Appli
8	680	9.0	2946	US-09-175-928-3	Sequence 2, Appli
9	267	3.5	1326	US-09-719-554-2	Sequence 20, Appl
10	211	2.8	711	US-09-719-554-20	Sequence 48, Appl
11	190	2.5	542	US-08-686-878A-48	Sequence 28, Appl
12	190	2.5	542	US-08-721-489-2	Sequence 28, Appl
13	183	2.4	1080	US-09-719-554-28	Sequence 21394, A
14	149	2.0	443	US-09-513-999C-21394	Sequence 21, Appl
15	89	1.2	711	US-09-719-554-21	Sequence 66, Appl
16	76	1.0	46340	US-09-719-554-66	Sequence 105, App
17	73	1.0	1481	US-08-979-847B-105	Sequence 13002, A
18	64	0.8	168394	US-09-949-016-13002	Sequence 5, Appli
19	63	0.8	1799	US-09-719-554-5	Sequence 15858, A
20	63	0.8	145320	US-09-949-016-15858	Sequence 112, App
21	58	0.8	758	US-08-979-847B-112	Sequence 12249, A
22	56	0.7	77772	US-09-949-016-17417	
23	56	0.7	77997	US-09-949-016-12249	

#### ALIGNMENTS

RESULT 1

55	0.7	2349	3	US-10-104-047-1148	Sequence 1148, Ap
55	0.7	149971	3	US-09-949-016-13590	Sequence 13590, A
54	0.7	1511	3	US-08-979-847B-117	Sequence 117, App
53	0.7	286	3	US-09-403-343B-32	Sequence 32, Appl
53	0.7	693	3	US-08-691-563C-88	Sequence 88, Appl
53	0.7	693	3	US-09-374-766-88	Sequence 88, Appl
53	0.7	693	3	US-08-979-847B-82	Sequence 82, Appl
53	0.7	1196	3	US-08-691-563C-56	Sequence 56, Appl
53	0.7	1196	3	US-09-374-766-56	Sequence 56, Appl
53	0.7	1196	3	US-08-979-847B-52	Sequence 52, Appl
53	0.7	1216	3	US-09-719-554-7	Sequence 7, Appli
53	0.7	1577	3	US-08-691-563C-89	Sequence 89, Appl
53	0.7	1577	3	US-09-374-766-89	Sequence 89, Appl
53	0.7	1577	3	US-08-979-847B-83	Sequence 83, Appl
53	0.7	2784	3	US-09-719-554-4	Sequence 4, Appli
51	0.7	601	3	US-09-949-016-202253	Sequence 202253,
50	0.7	279	2	US-08-686-878A-50	Sequence 50, Appl
50	0.7	279	2	US-08-721-489-4	Sequence 4, Appli
49	0.6	601	3	US-09-949-016-133739	Sequence 133739,
49	0.6	635	3	US-08-979-847B-102	Sequence 102, App
49	0.6	1774	3	US-09-719-554-17	Sequence 17, Appl
45	0.6	245286	3	US-09-949-016-15497	Sequence 15497, A
47	0.6	773	3	US-09-719-554-67	Sequence 67, Appl
47	0.6	1489	3	US-09-719-554-6	Sequence 6, Appli
46	0.6	601	3	US-09-949-016-202370	Sequence 202370,
46	0.6	84571	3	US-09-949-016-17420	Sequence 17420, A
44	0.6	764	3	US-08-979-847B-126	Sequence 126, App
44	0.6	800	3	US-08-979-847B-127	Sequence 127, App
44	0.6	1600	3	US-08-979-847B-206	Sequence 206, App
44	0.6	1600	3	US-08-979-847B-207	Sequence 207, App
43	0.6	506	3	US-09-621-976-2599	Sequence 2599, Ap
43	0.6	601	3	US-09-949-016-44297	Sequence 44297, A
43	0.6	601	3	US-09-949-016-202252	Sequence 202252,
43	0.6	1740	3	US-09-719-554-61	Sequence 61, Appl
43	0.6	2364	3	US-08-979-847B-88	Sequence 88, Appl
43	0.6	2391	3	US-08-691-563C-57	Sequence 57, Appl
43	0.6	2391	3	US-09-374-766-57	Sequence 57, Appl
43	0.6	2391	3	US-08-979-847B-53	Sequence 53, Appl
43	0.6	61108	3	US-09-949-002-608	Sequence 608, App
41	0.5	601	3	US-09-949-016-44296	Sequence 44296, A
41	0.5	601	3	US-09-949-016-202250	Sequence 202250,
41	0.5	601	3	US-09-949-016-202251	Sequence 202251,
41	0.5	1329	3	US-08-979-847B-108	Sequence 108, App
40	0.5	601	3	US-09-949-016-146560	Sequence 146560,
40	0.5	601	3	US-09-949-016-146561	Sequence 146561,
40	0.5	601	3	US-09-949-016-146564	Sequence 146564,
40	0.5	601	3	US-09-949-016-146565	Sequence 146565,
40	0.5	780	3	US-09-573-080A-385	Sequence 385, App
40	0.5	780	5	US-09-854-867-385	Sequence 205, App
40	0.5	1597	3	US-08-979-847B-205	Sequence 87, Appl
40	0.5	2304	3	US-08-979-847B-87	Sequence 44298, A
39	0.5	601	3	US-09-949-016-44298	Sequence 16, Appl
39	0.5	979	3	US-09-719-554-16	Sequence 195, App
38	0.5	85	4	US-08-979-847B-195	Sequence 538, App
38	0.5	410	4	US-08-880-107-538	Sequence 8, Appli
38	0.5	645	2	US-08-471-969-8	Sequence 8, Appli
38	0.5	645	2	US-08-384-137-8	Sequence 8, Appli
38	0.5	645	2	US-08-470-006A-8	Sequence 8, Appli
38	0.5	645	3	US-08-691-563C-8	Sequence 8, Appli
38	0.5	645	3	US-09-200-990-8	Sequence 8, Appli
38	0.5	645	3	US-09-133-411-8	Sequence 8, Appli
38	0.5	645	3	US-09-374-766-8	Sequence 8, Appli
38	0.5	645	3	US-08-979-847B-8	Sequence 8, Appli
38	0.5	1158	2	US-08-471-969-1	Sequence 1, Appli
38	0.5	1158	2	US-08-471-969-1	Sequence 1, Appli
38	0.5	1158	2	US-08-384-137-1	Sequence 1, Appli



```

US-09-719-554-1
; Sequence 1, Application US/09719554
; Patent No. 6919438
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
; FILE REFERENCE: 200936050PCT
; CURRENT APPLICATION NUMBER: US/09/719,554
; CURRENT FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2599
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-719-554-1

Query Match          10.3%; Score 779; DB 3; Length 2599;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 4802 AGGCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGCACCTAGTTCATGAAA 4861
DB 151 AGGCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGCACCTAGTTCATGAAA 210
QY 4862 TAAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACATAGCCCTGCTTCCAGG 4921
DB 211 TAAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACATAGCCCTGCTTCCAGG 270
QY 4922 CCACAGTAAACCAGGAGATCCAGGCGTTAGGTATACGATACATTTACACTGGCGCT 4981
DB 271 CCACAGTAAACCAGGAGTATCCAGGCGTTAGGTATACGATACATTTACACTGGCGCT 330
QY 4982 GAAGGCCACAGTCTCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAAGGACATCTAA 5041
DB 331 GAAGGCCACAGTCTCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAAAGGACATCTAA 390
QY 5042 AAAAGCAACCCAGGAAACCCACCTCACATGGCTGYCTGTGCTATAGCTTAAATA 5101
DB 391 AAAAGCAACCCAGGAAACCCACCTCACATGGCTGYCTGTGCTATAGCTTAAATA 450
QY 5102 GAATCTGCAACTTTCCTCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 5161
DB 451 GAATCTGCAACTTTCCTCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 510
QY 5162 CCTTCATAAACCAATGACCTTGTGTTGACCCAGAGCAGCAACTTAGTTGCGACATCAC 5221
DB 511 CCTTCATAAACCAATGACCTTGTGTTGACCCAGAGCAGCAACTTAGTTGCGACATCAC 570
QY 5222 CTCCTTAGCAATATCAACAGTCTTAAACATTTACAGGAACTATCCCTCGAGAGA 5281
DB 571 CTCCTTAGCAATATCAACAGTCTTAAACATTTACAGGAACTATCCCTCGAGAGA 630
QY 5282 GGGAAAAGAACTATTCCACCCWNGTGACATGGTATTAGTCAAGTCCCTCYCTCTAATTC 5341
DB 631 GGGAAAAGAACTATTCCACCCWNGTGACATGGTATTAGTCAAGTCCCTCYCTCTAATTC 690
QY 5342 CCCATCCCTAGATACATCTCGGGAAGGACCTACCCAGTCATTTTATYVACCCCAACTGC 5401
DB 691 CCCATCCCTAGATACATCTCGGGAAGGACCTACCCAGTCATTTTATCTACCCCAACTGC 750
QY 5402 GGTAAAGTGGCTGGAGTGGAGTCTGGATACATCACACTTGGTCAAAATCTGGATCT 5461
DB 751 GGTAAAGTGGCTGGAGTGGAGTCTGGATACATCACACTTGGTCAAAATCTGGATCT 810
QY 5462 GCCAAGGAACTTGAAAATCCAGGAGCAACGCTAGCTATTCTGTGAACCTCTAGAGGA 5521
DB 811 GCCAAGGAACTTGAAAATCCAGGAGCAACGCTAGCTATTCTGTGAACCTCTAGAGGA 870

```

```

QY 5522 TTTGGCCTGCTCTTCAAAACAACACGAGGAGAAAGTAACTAAATCATAAATCCCCCA 5581
DB 871 TTTGGCCTGCTCTTCAAAACAACACGAGGAGAAAGTAACTAAATCATAAATCCCCCA 929
QY 5582 TGGSCCTCCCTTATCATATTTTCTCTAGTGTSTTTTACCCCTSTTTCACTCTCACTG 5641
DB 930 TGGCCTCCCTTATCATATTTTCTCTTACTGTCTTTTACCCCTCTTTCACTCTCACTG 989
QY 5642 CACCCCTCCATGCGCTGTATGACCAAGTAGTCCCTTACCMAGAGTTTCTATGAGAA 5701
DB 990 CACCCCTCCATGCGCTGTATGACCAAGTAGTCCCTTACCAAGAGTTTCTATGAGAA 1049
QY 5702 TGCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGAAACCCCA 5761
DB 1050 TGCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTCTAAGGAAACCCCA 1109
QY 5762 CCTTCACTGCCACACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGC 5821
DB 1110 CCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGC 1169
QY 5822 ATGCAATATCTATTATTGGACAGGAAATGATTAATCTTAGTTGTCTTGGAGGACTTG 5881
DB 1170 ATGCAATATCTATTATTGGACAGGAAATGATTAATCTTAGTTGTCTTGGAGGACTTG 1229
QY 5882 GAGTCACTGTCTGTGGACTTACTTACCCAAAATGCTGTATGCTGATGGGGTGGAGTTC 5941
DB 1230 GAGTCACTGTCTGTGGACTTACTTACCCAAAATGCTGTATGCTGATGGGGTGGAGTTC 1289
QY 5942 AAGATCAGGCAAGAGAAAACATGTAAAGAAAGTAACTCTCCAACTCACCCGGGTACATG 6001
DB 1290 AAGATCAGGCAAGAGAAAACATGTAAAGAAAGTAACTCTCCAACTCACCCGGGTACATG 1349
QY 6002 GCACCTCTAGCCCTTCAAAAGGACTAGATCTCTCAAAACTACATGAAACCCCTCGTACCC 6061
DB 1350 GCACCTCTAGCCCTTCAAAAGGACTAGATCTCTCAAAACTACATGAAACCCCTCGTACCC 1409
QY 6062 ATACTCGCTCGTAAAGCCTATTATACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6121
DB 1410 ATACTCGCTCGTAAAGCCTATTATACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1469
QY 6122 AAAACCCCTACTAACTGTTGGATATGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6181
DB 1470 AAAACCCCTACTAACTGTTGGATATGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1529
QY 6182 CTGTACTCTGAAACAATGGAAACAATTCAGCAAGAAATAAAACCACTTCCGTTTGTAGTAG 6241
DB 1530 CTGTACTCTGAAACAATGGAAACAATTCAGCAAGAAATAAAACCACTTCCGTTTGTAGTAG 1589
QY 6242 GACCTCTTGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6301
DB 1590 GACCTCTTGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1649
QY 6302 GCAATCTACTACACAAACCACTCCCAATGCTCAGGTGGGTAACTCTCTCCCAACAA 6361
DB 1650 GCAATCTACTACACAAACCACTCCCAATGCTCAGGTGGGTAACTCTCTCCCAACAA 1709
QY 6362 TAGTCTGCTACCTCAGGAAATATTTTGTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6421
DB 1710 TAGTCTGCTACCTCAGGAAATATTTTGTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1769
QY 6422 ATGGCTCTTCAAGATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6481
DB 1770 ATGGCTCTTCAAGATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1829
QY 6482 CTGAAACAAGATTTATACAGTTATGCTATATCTTAAGCCCGCAACAAAAGAGTACCCATTC 6541
DB 1830 CTGAAACAAGATTTATACAGTTATGCTATATCTTAAGCCCGCAACAAAAGAGTACCCATTC 1889
QY 6542 TTCCCTTTTGTATAGGAGCAGAGTGTAGGTGCTAGGTACTGGCAATCTGGCAATCTGGCGGTATCA 6601
DB 1890 TTCCCTTTTGTATAGGAGCAGAGTGTAGGTGCTAGGTGCTAGGTACTGGCAATCTGGCGGTATCA 1949

```

QY 6602 CAACCTCTACTCAGTTCTTCAAACTATCTCAAGAACTAAATGGGACATGGAACGGG 6661  
 Db 1950 CAACCTCTACTCAGTTCTTCAAACTATCTCAAGAACTAAATGGGACATGGAACGGG 2009  
 QY 6662 TCGCCGACTCCCTGGTCACTTCCAGAGTCAAACTTAACTCCCTAGCAGCAGTAGTCTTTC 6721  
 Db 2010 TCGCCGACTCCCTGGTCACTTCCAGAGTCAAACTTAACTCCCTAGCAGCAGTAGTCTTTC 2069

RESULT 2  
 US-09-573-080A-21  
 ; Sequence 21, Application US/09573080A  
 ; Patent No. 6828097  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JOAN, KNOLL  
 ; APPLICANT: ROGAN, PETER  
 ; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATI  
 ; FILE REFERENCE: 30307  
 ; CURRENT APPLICATION NUMBER: US/09/573,080A  
 ; CURRENT FILING DATE: 2000-05-16  
 ; NUMBER OF SEQ ID NOS: 479  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 21  
 ; LENGTH: 8523  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: repeat region  
 ; LOCATION: (1)..(8523)  
 ; OTHER INFORMATION: herv17  
 ; PUBLICATION INFORMATION:  
 ; PUBLICATION INFORMATION:  
 ; AUTHORS: Jurka, J; Walichiewicz, J; Milosavljevic, A  
 ; TITLE: Prototypic sequences for human repetitive DNA  
 ; JOURNAL: Journal of Molecular Evolution  
 ; VOLUME: 35  
 ; ISSUE: 4  
 ; PAGES: 286-291  
 ; DATE: 1992-10-  
 ; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)  
 ; DATABASE ENTRY DATE: 1996-01-26  
 ; DATABASE ENTRY DATE: 1996-01-26  
 US-09-573-080A-21

Query Match 10.3%; Score 779; DB 3; Length 8523;  
 Best Local Similarity 98.9%; Pred. No. 0;  
 Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 4802 AGGCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATGAAGGCACTAGTTTCATGAAA 4861  
 Db 6071 AGGCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATGAAGGCACTAGTTTCATGAAA 6130  
 QY 4862 TAATTCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCAGG 4921  
 Db 6131 TAATTCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCAGG 6190  
 QY 4922 CCACAGTAAACCCAGGAGTATCCAGCGGTAGGTATACATATCACTTACATCGGCT 4981  
 Db 6191 CCACAGTAAACCCAGGAGTATCCAGCGGTAGGTATACATATCACTTACATCGGCT 6250  
 QY 4982 GAAGGCCACAGTCTCAGGAGAGGTGAGAAATGAATGAAYACTCAAAAGGACATCTAA 5041  
 Db 6251 GAAGGCCACAGTCTCAGGAGAGGTGAGAAATGAATGAAYACTCAAAAGGACATCTAA 6310  
 QY 5042 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCTGTCTGTGCTATAGCTTAAAAA 5101  
 Db 6311 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCTGTCTGTGCTATAGCTTAAAAA 6370  
 QY 5102 GAATCTGCAACTTCCCAAAAGAGGAGTCTAGCCCATACGAAATGCTGTATGGAAGGC 5161  
 Db 6371 GAATCTGCAACTTCCCAAAAGAGGAGTCTAGCCCATACGAAATGCTGTATGGAAGGC 6430  
 QY 5162 CCTTCATAACCAATGACCTTGTGCTTGAACCAAGACAGCCAACTTAGTTGTCAGACATCAC 5221

Db 6431 CTTTATATAACCAATGACCTTGTGCTTGAACCAAGACAGCCAACTTAGTTGACAGACATCAC 6490  
 QY 5222 CTCTTAGGCAAAATATCAAAAGTCTTAAACCATTTACAGGAACTATCTCCCTCGAGAAGA 5281  
 Db 6491 CTCTTAGGCAAAATATCAAAAGTCTTAAACCATTTACAGGAACTATCTCCCTCGAGAAGA 6550  
 QY 5282 GGGAAAAGAACTATTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTC 5341  
 Db 6551 GGGAAAAGAACTATTCCACCCCTTGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTC 6610  
 QY 5342 CCATCCCTAGATACATCTGGGAAGGACCTTACCAGTCACTTTATATATACCCCAACTGC 5401  
 Db 6611 CCATCCCTAGATACATCTGGGAAGGACCTTACCAGTCACTTTATATATACCCCAACTGC 6670  
 QY 5402 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAACTCTTGGATAC 5461  
 Db 6671 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAACTCTTGGATAC 6730  
 QY 5462 GCCAAAGGAACCTGAAAATCCAGGAGACAAACGCTAGCTATTCTGTGAACTCTAGAGGA 5521  
 Db 6731 GCCAAAGGAACCTGAAAATCCAGGAGACAAACGCTAGCTATTCTGTGAACTCTAGAGGA 6790  
 QY 5522 TTTGGCCTGCTCTTCAAAACAAACACAGGAGGAAAGTAAATCAATTAATCCCCCA 5581  
 Db 6791 TTTGGCCTGCTCTTCAAAACAAACACAGGAGGAAAGTAAATCAATTAATCCCCCA 6849  
 QY 5582 TGGSCCTCCCTTATCATATTTTCTCTKTAAGTGTSTTTTACCTCTTTTACCTCTCACTG 5641  
 Db 6850 TGGCCTCCCTTATCATATTTTCTCTTACTGTCTTTTACCTCTTTTACCTCTCACTG 6909  
 QY 5642 CACCCCTCCATGCGCTGTATGACAGTAGTCCCTTACCMAGAGTTTCTATAGAGAA 5701  
 Db 6910 CACCCCTCCATGCGCTGTATGACAGTAGTCCCTTACCAAGAGTTTCTATGAGAA 6969  
 QY 5702 TGCAGGCTCCCGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCA 5761  
 Db 6970 TGCAGGCTCCCGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCA 7029  
 QY 5762 CTTTCACTGCCCCACACCCATATGCCCCGCACTGCTATCACTCTGCACTCTTTTGTATGC 5821  
 Db 7030 CTTTCACTGCCCCACACCCATATGCCCCGCACTGCTATCACTCTGCACTCTTTTGTATGC 7089  
 QY 5822 ATGCAAAATACTCATTTATGGACAGGAAATAATGATTAATCTTAGTTCTCTGAGGAGCTTG 5881  
 Db 7090 ATGCAAAATACTCATTTATGGACAGGAAATAATGATTAATCTTAGTTCTCTGAGGAGCTTG 7149  
 QY 5882 GAGTCACTGTCTGTTGGACTTACTTCAACCAACTGGTATGCTGATGGGGTGGAGTTC 5941  
 Db 7150 GAGTCACTGTCTGTTGGACTTACTTCAACCAACTGGTATGCTGATGGGGTGGAGTTC 7209  
 QY 5942 AAGATCAGGCAAGAGAAAAACATGTAAAGAAAGTAACTCCCAACTCACCGGGTACATG 6001  
 Db 7210 AAGATCAGGCAAGAGAAAAACATGTAAAGAAAGTAACTCCCAACTCACCGGGTACATG 7269  
 QY 6002 GCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCCTCCGTACCC 6061  
 Db 7270 GCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCCTCCGTACCC 7329  
 QY 6062 ATACTGCTGTGTAGGCTTATTTAATACCACTCACTGGGCTCCATGAGGTCTCGGCC 6121  
 Db 7330 ATACTGCTGTGTAGGCTTATTTAATACCACTCACTGGGCTCCATGAGGTCTCGGCC 7389  
 QY 6122 AAAACCTCTACTTAACCTGTTGGATATGCTTCCCTGAACTTTCARGGCATATGTTCAATCC 6181  
 Db 7390 AAAACCTCTACTTAACCTGTTGGATATGCTTCCCTGAACTTTCARGGCATATGTTCAATCC 7449  
 QY 6182 CTGTACCTGAAACAAATGGAAACAACTTTCAGCAGAGAAATAAACACCACTTCCGTTTTAGTAG 6241  
 Db 7450 CTGTACCTGAAACAAATGGAAACAACTTTCAGCAGAGAAATAAACACCACTTCCGTTTTAGTAG 7509  
 QY 6242 GACCTCTTGTGTTTCCAAATSTGGAAATAACCCATACCTCAAACTCACTGTGTGTAATAATTTA 6301

Db 7510 GACCTCTTTTCCAATCTGGAAATAACCCATACCTCAAACTCACCTGTGTAAATTTA 7569

Qy 6302 GCAATACATACATACAAACCAACTCCCAATGATGATGAGTGGTAACTCTCTCCACACAAA 6361

Db 7570 GCAATACATACATACAAACCAACTCCCAATGATGATGAGTGGTAACTCTCTCCACACAAA 7629

Qy 6362 TAGTCTGCTACCTCAGGAAATATTTTGTCTGTGTGAGTACCTCAGCCATATCGTTGTTGA 6421

Db 7630 TAGTCTGCTACCTCAGGAAATATTTTGTCTGTGTGAGTACCTCAGCCATATCGTTGTTGA 7689

Qy 6422 ATGGCTCTTCAGAACTATGTGCTTCTCTCATTTCTTAGTGCCCCYATGRCAATCTACA 6481

Db 7690 ATGGCTCTTCAGAACTATGTGCTTCTCTCATTTCTTAGTGCCCCYATGRCAATCTACA 7749

Qy 6482 CTGAACAAGATTTATACAGTTATGTATCATATCTAAGCCCGGCAACAAAAGAGTACCATTTC 6541

Db 7750 CTGAACAAGATTTATACAGTTATGTATCATATCTAAGCCCGGCAACAAAAGAGTACCATTTC 7809

Qy 6542 TTCCCTTTTGTATAGGACGAGTGCTAGGTGCACTAGGTACTGGCAATTTGGCGGTATCA 6601

Db 7810 TTCCCTTTTGTATAGGACGAGTGCTAGGTGCACTAGGTACTGGCAATTTGGCGGTATCA 7869

Qy 6602 CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAAACGGG 6661

Db 7870 CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAAACGGG 7929

Qy 6662 TGGCGAGTCCCTGTGCTACCTTGGCAAGATCAACTTAACTCCCTAGCAGAGTAGTCTTTC 6721

Db 7930 TGGCGAGTCCCTGTGCTACCTTGGCAAGATCAACTTAACTCCCTAGCAGAGTAGTCTTTC 7989

RESULT 3

US-09-854-867-21

; Sequence 21, Application US/09854867

; Patent No. 7014997

; GENERAL INFORMATION:

; APPLICANT: JOAN, KNOLL H

; APPLICANT: ROGAN, PETER K

; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING

; FILE REFERENCE: 30307

; CURRENT APPLICATION NUMBER: US/09/854,867

; CURRENT FILING DATE: 2003-05-08

; NUMBER OF SEQ ID NOS: 613

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 21

; LENGTH: 8523

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: repeat\_region

; LOCATION: (1)..(8523)

; OTHER INFORMATION: herv17

US-09-854-867-21

Query Match 10.3%; Score 779; DB 5; Length 8523;

Best Local Similarity 98.9%; Pred. No. 0;

Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

Qy 4802 AGGCCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATGAAGCAGCTAGTTCAATGAAA 4861

Db 6071 AGGCCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATGAAGCAGCTAGTTCAATGAAA 6130

Qy 4862 TAAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCCAGG 4921

Db 6131 TAAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCCAGG 6190

Qy 4922 CCAAGTAAACCCAGGAGTATCCAGGCGTTAGGTATAGCATATCACTTACAGTGGCCT 4981

Db 6191 CCAAGTAAACCCAGGAGTATCCAGGCGTTAGGTATAGCATATCACTTACAGTGGCCT 6250

Qy 4982 GAAAGCCAGTCTCAGGGAAGTTCGAGAAAATGAATGAAAYACTCAAGAGGACATCTAA 5041

Db 6251 GAAAGCCAGTCTCAGGGAAGTTCGAGAAAATGAATGAAAYACTCAAGAGGACATCTAA 6310

Qy 5042 AARAGCAAAACCCAGGAACCCCACTCAATGGCCTGTCTGTGCTATAGCTTTAAAAA 5101

Db 6311 AARAGCAAAACCCAGGAACCCCACTCAATGGCCTGTCTGTGCTATAGCTTTAAAAA 6370

Qy 5102 GAATCTGCAACTTTTCCCAAAAGCAGGACTTAGCCCATAGCAAAATGCTGTATGGAAGGC 5161

Db 6371 GAATCTGCAACTTTTCCCAAAAGCAGGACTTAGCCCATAGCAAAATGCTGTATGGAAGGC 6430

Qy 5162 CTTTCATAACCAATGACCTTGTGTGACCCAAAGACGCCAACTTAGTTGCAAGACATCAC 5221

Db 6431 CTTTCATAACCAATGACCTTGTGTGACCCAAAGACGCCAACTTAGTTGCAAGACATCAC 6490

Qy 5222 CTCCTTAGCCAAATATCAAGAGTTCTTAAACATTACAAGAACCTTATCCCTGGAAGA 5281

Db 6491 CTCCTTAGCCAAATATCAAGAGTTCTTAAACATTACAAGAACCTTATCCCTGGAAGA 6550

Qy 5282 GGGAAAAGAACTATTTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCTCTAAATTC 5341

Db 6551 GGGAAAAGAACTATTTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCTCTAAATTC 6610

Qy 5342 CCCATCCCTAGATACATCTCTGGGAAGGACCTA CCCAGTCAATTTATYTAACCCCAACTGC 5401

Db 6611 CCCATCCCTAGATACATCTCTGGGAAGGACCTA CCCAGTCAATTTATCTACCCCAACTGC 6670

Qy 5402 GGTAAAAGTGGCTGGAGTGTCTTGGATACATCACATTTAGTCAAAATCCTTGATACT 5461

Db 6671 GGTAAAAGTGGCTGGAGTGTCTTGGATACATCACATTTAGTCAAAATCCTTGATACT 6730

Qy 5462 GCCAAAAGAACTGAAAATCCAGAGACAAAGCTAGCTATTCTCTGTAACCTCTAGAGGA 5521

Db 6731 GCCAAAAGAACTGAAAATCCAGAGACAAAGCTAGCTATTCTCTGTAACCTCTAGAGGA 6790

Qy 5522 TTTGGCGCTGCTCTTCAAAACAAACCCAGGAGGAAAGTAACTAAATCATAAATCCCCCA 5581

Db 6791 TTTGGCGCTGCTCTTCAAAACAAACCCAGGAGGAAAGTAACTAAATCATAAATCCCCCA 6849

Qy 5582 TGGSGCTCCCTTATCATATTTTCTCTATAGTCTTCTTATAGCTCTTCTCTCTCTCTCT 5641

Db 6850 TGGSGCTCCCTTATCATATTTTCTCTATAGTCTTCTTACTCTTCTTCTCTCTCTCTCT 6909

Qy 5642 CACCCCTCCATGCGCTGTATGACAGTCTCCCTCTCCCTCTTCTTATGAGGAA 5701

Db 6910 CACCCCTCCATGCGCTGTATGACAGTCTCCCTCTCCCTCTTCTTATGAGGAA 6969

Qy 5702 TGCAAGCTCCCGGAAATATGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCA 5761

Db 6970 TGCAAGCTCCCGGAAATATGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCA 7029

Qy 5762 CTTTCACTGCCCCACACCATATGCCCCGCAACTGCTATCACTGCGCACTCTTTCATGCG 5821

Db 7030 CTTTCACTGCCCCACACCATATGCCCCGCAACTGCTATCACTGCGCACTCTTTCATGCG 7089

Qy 5822 ATGCAATACTCATTTTGGACAGAAAATGATTAATCTCTAGTTGTCTCTGAGGACTTG 5881

Db 7090 ATGCAATACTCATTTTGGACAGAAAATGATTAATCTCTAGTTGTCTCTGAGGACTTG 7149

Qy 5882 GAGTCACTGTCTGTTGGACTTATTTACCCAAACTGGTATGTCTCATGAGGAGTCTGAGTTC 5941

Db 7150 GAGTCACTGTCTGTTGGACTTATTTACCCAAACTGGTATGTCTCATGAGGAGTCTGAGTTC 7209

Qy 5942 AAGTCAAGGAGAGAAAACATGTAAGAAAGTAACTCTCCCACTCACCSSGGGTACATG 6001

Db 7210 AAGTCAAGGAGAGAAAACATGTAAGAAAGTAACTCTCCCACTCACCSSGGGTACATG 7269

Qy 6002 GCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAAACATCATGAAACCTCCGTAACCC 6061

Db 7270 GCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAAACATCATGAAACCTCCGTAACCC 7329

Qy 6062 ATACTCGCTGGTAAAGCCTATTTAATACCAACCTCACTGGGCTCCTAGAGTCTCGGCC 6121

Db 7330 ATACTCGCTGGTAAAGCCTATTTAATACCAACCTCACTGGGCTCCTAGAGTCTCGGCC 7389







Db	9717	CTTCACTGCCACACCCATATGCGCGCAACTGCTATCACTCTGCCACTCTTTGTCATGC	9776
Qy	5822	ATCCAAATACCTATTATGACAGGAAAAATGATTAATCTAGTTGCTCGAGGACTTG	5881
Db	9777	ATGCAATACCTATTATGACAGGAAAAATGATTAATCTAGTTGCTCGAGGACTTG	9836
Qy	5882	GAGTCACCTGCTCTGTGGACTTACTTACCCCAAACTGCTATGCTGATGGGGTGGAGTTTC	5941
Db	9837	GAGTCACCTGCTCTGTGGACTTACTTACCCCAAACTGCTATGCTGATGGGGTGGAGTTTC	9896
Qy	5942	AGATCAGGCAAGAGAAAAACATGTAAAGAACTAATCTCCCAACTCACCGGGTACATG	6001
Db	9897	AGATCAGGCAAGAGAAAAACATGTAAAGAACTAATCTCCCAACTCACCGGGTACATG	9956
Qy	6002	GCACCTCTAGCCCTCAAGAGACTAGATCTCTCAAAACTACATGATGAACCCCTCGTACCC	6061
Db	9957	GCACCTCTAGCCCTCAAGAGACTAGATCTCTCAAAACTACATGATGAACCCCTCGTACCC	10016
Qy	6062	ATACTGCGCTGGTAAGCCTATTATTAACACACCTCACTGGGCTCCATGAGTCTCGGCC	6121
Db	10017	ATACTGCGCTGGTAAGCCTATTATTAACACACCTCACTGGGCTCCATGAGTCTCGGCC	10076
Qy	6122	AAACCCCTACTAATCTTGGATATGCTCCCTCGAACTTCARGCCATATGTTTCAATCC	6181
Db	10077	AAACCCCTACTAATCTTGGATATGCTCCCTCGAACTTCARGCCATATGTTTCAATCC	10136
Qy	6182	CTGTACTGAACAAATGGAACTTTCAGACAGAAATAAACCACCTCGTTTGTAGTAG	6241
Db	10137	CTGTACTGAACAAATGGAACTTTCAGACAGAAATAAACCACCTCGTTTGTAGTAG	10196
Qy	6242	GACCTCTGTTTCCAAATGGAATAAACCATACCTCAACCTCACTGTTGTAATAATTTA	6301
Db	10197	GACCTCTGTTTCCAAATGGAATAAACCATACCTCAACCTCACTGTTGTAATAATTTA	10256
Qy	6302	GCAATACTACATACACCACTCCCAATGCACTCAGGTGGGTAACTCTCTCCACACAAA	6361
Db	10257	GCAATACTACATACACCACTCCCAATGCACTCAGGTGGGTAACTCTCTCCACACAAA	10316
Qy	6362	TAGTCTGCTACCTCAGGAATATTTTGTCTGCGGTACCTCAGCCTATCGTTGTTGA	6421
Db	10317	TAGTCTGCTACCTCAGGAATATTTTGTCTGCGGTACCTCAGCCTATCGTTGTTGA	10376
Qy	6422	ATGGCTCTTCAGAACTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	6481
Db	10377	ATGGCTCTTCAGAACTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	10436
Qy	6482	CTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCGCAACAAAGAGTACCCATTC	6541
Db	10437	CTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCGCAACAAAGAGTACCCATTC	10496
Qy	6542	TTCCCTTTTGTATAGGAGCAGAGTGTAGGTGCACTAGGTACTGCGCATTTGGCGGTATCA	6601
Db	10497	TTCCCTTTTGTATAGGAGCAGAGTGTAGGTGCACTAGGTACTGCGCATTTGGCGGTATCA	10556
Qy	6602	CAACCTCTACTAGTTCTACTACAACTATCTCAAGAACTAAATGGGGCATGGAACGGG	6661
Db	10557	CAACCTCTACTAGTTCTACTACAACTATCTCAAGAACTAAATGGGGCATGGAACGGG	10616
Qy	6662	TGCGCGACTCCCTGGTCACTTGCAGAGATCACTTAACCTCCCTAGCAGAGTACCTTTC	6721
Db	10617	TGCGCGACTCCCTGGTCACTTGCAGAGATCACTTAACCTCCCTAGCAGAGTACCTTTC	10676

RESULT 6

US-09-719-554-22  
; Sequence 22, Application US/09719554  
; Patent No. 6919438  
; GENERAL INFORMATION:  
; APPLICANT: ALLIEL, Patrick  
; APPLICANT: PERIN, Jean-Pierre  
; APPLICANT: RIEGER, Francois  
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU  
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS

FILE REFERENCE: 200936USOPCT  
; CURRENT APPLICATION NUMBER: US/09/719,554  
; CURRENT FILING DATE: 2001-12-26  
; PRIOR APPLICATION NUMBER: PCT/FR99/01513  
; PRIOR FILING DATE: 1999-06-25  
; NUMBER OF SEQ ID NOS: 122  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 22  
; LENGTH: 2055  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(2055)  
; OTHER INFORMATION:  
US-09-719-554-22

Query Match		9.6%;	Score 731;	DB 3;	Length 2055;
Best Local Similarity		99.4%;	Pred. No. 0;		
Matches 1031;		Conservative 0;	Mismatches 6;	Indels 0;	Gaps 0;
Qy	5685	AGAGTTTCTTATGGAGAATGCAGCGTCCCGGAAATATTTGATGCCCATCGTATAGGAGTCT	5744		
Db	495	AGAGTTTCTTATGGAGAATGCAGCGTCCCGGAAATATTTGATGCCCATCGTATAGGAGTCT	554		
Qy	5745	TTSTAAGGGAACCCCACTTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTC	5804		
Db	555	TTCTAAGGGAACCCCACTTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTC	614		
Qy	5805	TGCCACTCTTTTCATGCAATGCAATCACTTATTTGGACAGGAAATGATTAATCTTAG	5864		
Db	615	TGCCACTCTTTTCATGCAATGCAATCACTTATTTGGACAGGAAATGATTAATCTTAG	674		
Qy	5865	TTGTCTCTGAGGACTTTGGAGTCACTGTCTGTGGACTTACTTCAACCAACTGGTATGTC	5924		
Db	675	TTGTCTCTGAGGACTTTGGAGTCACTGTCTGTGGACTTACTTCAACCAACTGGTATGTC	734		
Qy	5925	TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAAAGTAAATCTCCCA	5984		
Db	735	TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAAAGTAAATCTCCCA	794		
Qy	5985	ACTCAGCGGGTACATGGGACCTCTAGCCCTCAAGAGACTAGATCTCTCAAAACTACA	6044		
Db	795	ACTCAGCGGGTACATGGGACCTCTAGCCCTCAAGAGACTAGATCTCTCAAAACTACA	854		
Qy	6045	TGAAACCTCTCGTACCCATCTCGGCTGCTAAGCCTATTTAATACCACTCCTCACTGGGT	6104		
Db	855	TGAAACCTCTCGTACCCATCTCGGCTGCTAAGCCTATTTAATACCACTCCTCACTGGGT	914		
Qy	6105	CCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTTGGATATGCTCCCTCCCTGAACTTCAR	6164		
Db	915	CCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTTGGATATGCTCCCTCCCTGAACTTCAR	974		
Qy	6165	GCCATATGTTTCAATCCCTGTACCTGAAACAATTTGAAACAATTTGAGCAAGATAAACAC	6224		
Db	975	GCCATATGTTTCAATCCCTGTACCTGAAACAATTTGAAACAATTTGAGCAAGATAAACAC	1034		
Qy	6225	CACCTTCGTTTGTAGTAGGACCTCTGTTTCCAAATSTGGAAATAACCCATACCTCAAACT	6284		
Db	1035	CACCTTCGTTTGTAGTAGGACCTCTGTTTCCAAATSTGGAAATAACCCATACCTCAAACT	1094		
Qy	6285	CACCTGTGTAAATTTTAGCAATATCTACATACCAACCAACTCCCAATGATCAGGTGGGT	6344		
Db	1095	CACCTGTGTAAATTTTAGCAATATCTACATACCAACCAACTCCCAATGATCAGGTGGGT	1154		
Qy	6345	AACCTCTCCCAACAATAGTCTGCTACCTCAGGAATATTTTGTGTGTGTTGTTGTTGTTG	6404		
Db	1155	AACCTCTCCCAACAATAGTCTGCTACCTCAGGAATATTTTGTGTGTGTTGTTGTTGTTG	1214		
Qy	6405	AGCCTATCGTTGTTTGAATGCTCTTCAAGATCTATGCTCTCTCTCTCTCTCTCTCTCTCT	6464		
Db	1215	AGCCTATCGTTGTTTGAATGCTCTTCAAGATCTATGCTCTCTCTCTCTCTCTCTCTCTCT	1274		



; ORGANISM: Homo sapiens  
US-09-175-928-3

```

Query Match      9.0%; SCORE 680; DB 3; Length 2946;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1030; Conservative 0; Mismatches 7; Indels

```

5685	QY	AGAGTTTCTTATGAGGAATGACGGTCCCGGAAATATTGATGCCCCATCGTATAGAGTCT	5744
1032	Db	AGAGTTTCTATGAGGAATGACGGTCCCGGAAATATTGATGCCCCATCGTATAGAGTCT	1091
5745	QY	TTSTAAGGGAAACCCCACTTCACTGCGCCACACCATATGCCCCGAACTGCTATCACTC	5804
1092	Db	TTCTAAGGGAAACCCCACTTCACTGCGCCACCCATATGCCCCGAACTGCTATCACTC	1151
5805	QY	TGCCACTCTTTGCAATGCATGCAAAATCTCAATTATGGACAGGAAAAATGATTATCCTAG	5864
1152	Db	TGCCCACTCTTTGCAATGCATGCAAAATCTCAATTATGGACAGGAAAAATGATTATCCTAG	1211
5865	QY	TTCTCTCTGGAGGACTTTGGAGTCACTGCTCTGTGGACCTTATCTCACCAAACTGGTATGTC	5924
1212	Db	TTGTCTCTGGAGGACTTTGGAGTCACTGCTCTGTGGACCTTATCTCACCAAACTGGTATGTC	1271
5925	QY	TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCA	5984
1272	Db	TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCA	1331
5985	QY	ACTCACSGGGTACATGGCACTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACA	6044
1332	Db	ACTCACCGGGTACATGGCACTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACA	1391
6045	QY	TGAACCCCTCCGTACCCATACCTCGCTGGTAAAGCTATTTTAATACCAACCCTCACTGGGCT	6104
1392	Db	TGAAACCCCTCCGTACCCATACCTCGCTGGTAAAGCTATTTTAATACCAACCCTCACTGGGCT	1451
6105	QY	CAATGAGGTCTCGGCCCAAAACCCTACTAATCTGTGGATATGCTCCCCCTGAACCTTCAR	6164
1452	Db	CAATGAGGTCTCGGCCCAAAACCCTACTAATCTGTGGATATGCTCCCCCTGAACCTTCAG	1511
6165	QY	GCCATATGTTTCAATCCCTGTACTCGAACAAATGGAACTTCAGACACAGAAATAAACAC	6224
1512	Db	GCCATATGTTTCAATCCCTGTACTCGAACAAATGGAACTTCAGACACAGAAATAAACAC	1571
6225	QY	CACCTCCGTTTATGATAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACT	6284
1572	Db	CACCTCCGTTTATGATAGGACCTCTTGTTTCCAATSTGGAAATAACCCATACCTCAAACT	1631
6285	QY	CACCTGTGTAAATTTAGCAATACTACATACAAACCACTCCCAATGCATCAGTGGGT	6344
1632	Db	CACCTGTGTAAATTTAGCAATACTACATACAAACCACTCCCAATGCATCAGTGGGT	1691
6345	QY	AACTCTCCCAACAATAATAGTCTGCCCTACCCCTCAGGAATATTTTTGTCTGTGGTACCTC	6404
1692	Db	AACTCTCCCAACAATAATAGTCTGCCCTACCCCTCAGGAATATTTTTGTCTGTGGTACCTC	1751
6405	QY	AGCCTATCGTTGTTTGAATGGCTCTTCAGAACTATATGTGCTTCCCTCATCTTTAGTGCC	6464
1752	Db	AGCCTATCGTTGTTTGAATGGCTCTTCAGAACTATATGTGCTTCCCTCATCTTTAGTGCC	1811
6465	QY	CCCYATGRCCTACTACACTGAACAAGATTTATACAGTTATGTCAATCTAAGCCCGCAAA	6524
1812	Db	CCCTATGACCACTACACTGAACAAGATTTATACAGTTATGTCAATCTAAGCCCGCAAA	1871
6525	QY	CAAAAGATACCAATCTTCTCTTTGTTATAGAGCAGGAGTGTAGGTGCACTAGGTAC	6584
1872	Db	CAAAAGATACCAATCTTCTCTTTGTTATAGAGCAGGAGTGTAGGTGCACTAGGTAC	1931
6585	QY	TGGCATTTGGCGGTATCAACACTCTACTCAGTTCTACTACAACTACTCAAGAACTAAA	6644
1932	Db	TGGCATTTGGCGGTATCAACACTCTACTCAGTTCTACTACAACTACTCAAGAACTAAA	1991
6645	QY	TGGGGAATGGAAACGGGTCCGCCACTCCCTGGTCACTTTGCAAGATCAAACTTAACCTCCCT	6704

Db 1992 TGGGGA CATGGAACGGGTCCCGACTCCTTGGTCACTTCAAGATCAACTTAATCCCT 2051

Qy 6705 AGCAGCAGTAGTCCTTC 6721  
|||||

Db 2052 AGCAGCAGTAGTCCTTC 2068

RESIT.T 9

```

US-09-719-554-2
; Sequence 2, Application US/09719554
; Patent No. 6919438
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL
; FILE REFERENCE: 20093650PCT
; CURRENT APPLICATION NUMBER: US/09/719,554
; CURRENT FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1326
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-719-554-2

```

Query Match 3.5%; Score 267; DB 3; Length 1326;  
Best Local Similarity 99.4%; Pred. No. 3.8e-124;

Matches	487; Conservative	0; Mismatches	2; Indels	1; Gaps	1;
QY	1507	GGGAAGGGAAGAAATCTCTACTGCGCTTTCTGGAGAGACTAAAGGAGGCATTGAGGAAGCG	1566		
Db	779	GGGAAGGGAAGAAATCTCTACTGCGCTTTCTGGAGAGACTAAAGGAGGCATTGAGGAAGCG	838		
QY	1567	TGCGCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCAC	1626		
Db	839	TGCGCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTTATCAC	898		
QY	1627	TCAGTCAGCTGCAGACATTAG-AAAAAATCTCAAAAGTCTGCCGTAGGCCCGCAGAGCAAAA	1685		
Db	899	TCAGTCAGCTGCAGACATTAGAAAAAATCTCAAAAGTCTGCCGTAGGCCCGCAGAGCAAAA	958		
QY	1686	CTTAGAAACCTATTGAACCTTGCGCAACYTCGGTTTTTATAATAGAGATCAGAGGAGCA	1745		
Db	959	CTTAGAAACCTATTGAACCTTGCGCAACCTCGGTTTTTATAATAGAGATCAGAGGAGCA	1018		
QY	1746	GGCGGAACAGGACAAACCGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCCCTCAGGC	1805		
Db	1019	GGCGGAACAGGACAAACCGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCCCTCAGGC	1078		
QY	1806	AAGTGGACCTTTCGAGGCTCTGGAAAGGGAAGCTGGGCAAAATGCAATGCCTTAATAGG	1865		
Db	1079	AAGTGGACCTTTCGAGGCTCTGGAAAGGGAAGCTGGGCAAAATGCAATGCCTTAATAGG	1138		
QY	1866	CTTGCTTCCAGTCGGCTCTACAAGGACACTTTAAAAAGATTGTCACAGTAGAAGTAAGC	1925		
Db	1139	CTTGCTTCCAGTCGGCTCTACAAGGACACTTTAAAAAGATTGTCACAGTAGAAGTAAGC	1198		
QY	1926	CGCCCCCTTCGTCATGCCCTTTATTTCAAGGGAATCAGCTGGAAAGGCCCACTGCCCCAGGG	1985		
Db	1199	CGCCCCCTTCGTCATGCCCTTTATTTCAAGGGAATCAGCTGGAAAGGCCCACTGCCCCAGGG	1258		
QY	1986	GACAAAGGTC	1995		
Db	1259	GACAAAGGTC	1268		

RESULT 10  
US-09-719-554-20

```
; Sequence 20, Application US/09719554
; Patent No. 6919438
; GENERAL INFORMATION:
; APPLICANT: ALLIED, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
; FILE REFERENCE: 200936US0PCT
; CURRENT APPLICATION NUMBER: US/09/719,554
; CURRENT FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; TYPE: DNA
; LENGTH: 711
; ORGANISM: Homo sapiens
US-09-719-554-20

Query Match      2.8%; Score 211; DB 3; Length 711;
Best Local Similarity 99.6%; Pred. No. 8e-96;
Matches 261; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCC 60
DB 209 CAACATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCC 268

QY 61 CTTCCCTTTGTATGGGAGCTGTTTTTCATGCTATTTTCACTCTATTAAATCTTGCAACTGCA 120
DB 269 CTTCCCTTTGTATGGGAGCTGTTTTTCATGCTATTTTCACTCTATTAAATCTTGCAACTGCA 328

QY 121 CTCCTTGTGCTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTCACCGTCCACCACCTGC 180
DB 329 CTCCTTGTGCTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTCACCGTCCACCACCTGC 388

QY 181 TGTTTGCCACACCGCCANACCTCCGCTGACTCCCATCCCTCTGGATCTGCGAGGTGC 240
DB 389 TGTTTGCCACACCGCCAGACCTGCGCTGACTCCCATCCCTCTGGATCTGCGAGGTGC 448

QY 241 CGCTGTGCTCCTGATCCAGCGA 262
DB 449 CGCTGTGCTCCTGATCCAGCGA 470

RESULT 11
US-08-686-878A-48
; Sequence 48, Application US/08686878A
; Patent No. 5708157
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: LaVallie, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Evans, Cheryl
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn version 3.1
; CURRENT APPLICATION NUMBER: US/08/721,489
; APPLICATION FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/686,878A
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 542 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-686-878A-48

Query Match      2.5%; Score 190; DB 2; Length 542;
Best Local Similarity 100.0%; Pred. No. 3.3e-85;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5390 TACCCCAACTCGGTTAAAGTGGCTGGAGTGGAGTCTTTGGATACATCACACTTGAGTCAA 5449
DB 152 TACCCCAACTCGGTTAAAGTGGCTGGAGTGGAGTCTTTGGATACATCACACTTGAGTCAA 211

QY 5450 ATCTCTGATATCTGCCAAAGAACCTTGAAATCCAGGAGACAAACGCTAGCTATTCCTGTGA 5509
DB 212 ATCTCTGATATCTGCCAAAGAACCTTGAAATCCAGGAGACAAACGCTAGCTATTCCTGTGA 271

QY 5510 ACCTCTAGAGATTGGCCCTGCTCTTCAAAACAACACAGAGAGAAAGTAACCTAAATC 5569
DB 272 ACCTCTAGAGATTGGCCCTGCTCTTCAAAACAACACAGAGAGAAAGTAACCTAAATC 331

QY 5570 ATAAATCCCC 5579
DB 332 ATAAATCCCC 341

RESULT 12
US-08-721-489-2
; Sequence 2, Application US/08721489
; Patent No. 5786465
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: LaVallie, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/721,489
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
```

```

; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 542 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-721-489-2

Query Match      2.5%; Score 190; DB 2; Length 542;
Best Local Similarity 100.0%; Pred. No. 3.3e-85;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5390 TACCCCAACTGCGGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAA 5449
Db 152 TACCCCAACTGCGGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAA 211

QY 5450 ATCCTGGATCTACCCAAAGGAACCTGAAAATCCAGGAGACAAACGCTAGCTATTTCCTGTGA 5509
Db 212 ATCCTGGATCTACCCAAAGGAACCTGAAAATCCAGGAGACAAACGCTAGCTATTTCCTGTGA 271

QY 5510 ACCTCTAGAGGATTGGCGCTGCTCTTCAAAACCAACAGGAGGAAAGTAACTAAATC 5569
Db 272 ACCTCTAGAGGATTGGCGCTGCTCTTCAAAACCAACAGGAGGAAAGTAACTAAATC 331

QY 5570 ATAAATCCCC 5579
Db 332 ATAAATCCCC 341

RESULT 13
US-09-719-554-28
; Sequence 28, Application US/09719554
; Patent No. 6919438
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
; FILE REFERENCE: 200936USOPT
; CURRENT APPLICATION NUMBER: US/09/719,554
; PRIOR FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 28
; LENGTH: 1080
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1080)
; OTHER INFORMATION:
US-09-719-554-28

Query Match      2.4%; Score 183; DB 3; Length 1080;
Best Local Similarity 99.4%; Pred. No. 1.2e-81;
Matches 353; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1507 GGGAGGGAGAAAATCTCTACTGCCCTTTCTGGAGAGACTAAAGGAGGCATTGAGGAAGCG 1566
Db 726 GGGAGGGAGAAAATCTCTACTGCCCTTTCTGGAGAGACTAAAGGAGGCATTGAGGAAGCG 785

QY 1567 TGCCTCTCTGTACCTGACTCTCTTGAAGCCCACTAACTTAAAGCGTAAGTTTATCAC 1626
Db 786 TGCCTCTCTGTACCTGACTCTCTTGAAGCCCACTAACTTAAAGCGTAAGTTTATCAC 845

```

```

QY 1627 TCAGTCAGCTGCAGACATTTAG-AAAAAATTTCAAAAGTCTGCCGTAGGCCGAGCAAAA 1685
Db 846 TCAGTCAGCTGCAGACATTTAGAAAAAATTTCAAAAGTCTGCCGTAGGCCGAGCAAAA 905

QY 1686 CTTAGAAAACCTATTCAACTTGGCAACCTCGGTTTATTAATAGATAGATAGAGAGGAGCA 1745
Db 906 CTTAGAAAACCTATTCAACTTGGCAACCTCGGTTTATTAATAGATAGATAGAGAGGAGCA 965

QY 1746 GCGGAAACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCTCAGGC 1805
Db 966 GCGGAAACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCTCAGGC 1025

QY 1806 AAGTGAAGCTTTGAGGCTCTGGAAAAGGGAAGCTGGGCAAAATTTGAATGCCTAA 1860
Db 1026 AAGTGAAGCTTTGAGGCTCTGGAAAAGGGAAGCTGGGCAAAATTTGAATGCCTAA 1080

```

```

RESULT 14
US-09-513-999C-21394
; Sequence 21394, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expresed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pn
; SEQ ID NO 21394
; LENGTH: 443
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 102
; OTHER INFORMATION: k=g or t
US-09-513-999C-21394

```

```

Query Match      2.0%; Score 149; DB 3; Length 443;
Best Local Similarity 99.2%; Pred. No. 1.8e-64;
Matches 249; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 12 ATATAAACCCAGGCAATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCCCTTCCCTTTGT 71
Db 3 ATATAAACCCAGGCAATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCCCTTCCCTTTGT 62

QY 72 ATGGGAGCTGTTTTCATGCTATTTCACTCTATTAATCTTGAACATGCACTCTCTGGTC 131
Db 63 ATGGGAGCTGTTTTCATGCTATTTCACTCTATTAATCTTGAACATGCACTCTCTGGTC 122

QY 132 CATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTCCAGTCCACCACTGCTGTTGCGACC 191
Db 123 CATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTCCAGTCCACCACTGCTGTTGCGACC 182

QY 192 ACCGCANACCTGCCGCTGACTCCCATCCCTCTGGATCCCTGAGGGTGTCCGCTGTGCTCC 251
Db 183 ACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCCTGAGGGTGTCCGCTGTGCTCC 242

QY 252 TCATCCAGCA 262
Db 243 TCATCCAGCA 253

RESULT 15
US-09-719-554-21
; Sequence 21, Application US/09719554

```

```

; Patent No. 6919438
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS
; FILE REFERENCE: 200936US0PCT
; CURRENT APPLICATION NUMBER: US/09/719,554
; CURRENT FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-719-554-21

```

```

Query Match 1.2%; Score 89; DB 3; Length 711;
Best Local Similarity 100.0%; Pred. No. 4.1e-34;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7488 TTCGAGCGGCAACGCAACCCCTTGGGTCCCTCCCTTGTATGGAGCTCTGTTT 7547
Db 235 TTCGAGCGGCAACGCAACCCCTTGGGTCCCTCCCTTGTATGGAGCTCTGTTT 294

QY 7548 CATGCTATTTCACCTATTAAATCTTGCA 7576
Db 295 CATGCTATTTCACCTATTAAATCTTGCA 323

```

```

RESULT 16
US-09-719-554-66
; Sequence 66, Application US/09719554
; Patent No. 6919438
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS
; FILE REFERENCE: 200936US0PCT
; CURRENT APPLICATION NUMBER: US/09/719,554
; CURRENT FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66
; LENGTH: 46340
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-719-554-66

```

```

Query Match 1.0%; Score 76; DB 3; Length 46340;
Best Local Similarity 100.0%; Pred. No. 1.7e-27;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6068 GCCTGGTAGCCTATTATACACCCCTCAGTGGGCTCCATGAGTCTCGGCCCAAAACC 6127
Db 32342 GCCTGGTAGCCTATTATACACCCCTCAGTGGGCTCCATGAGTCTCGGCCCAAAACC 32401

QY 6128 CTACTAACTGTTGGAT 6143
Db 32402 CTACTAACTGTTGGAT 32417

```

```

RESULT 17
US-08-979-847B-105
; Sequence 105, Application US/08979847B
; Patent No. 6582703

```

```

; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; APPLICANT: BESEME, FREDERIC
; APPLICANT: BEDIN, FREDERIC
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: KOMURIAN-PRADEL, FLORENCE
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: MANDRAND, BERNARD
; APPLICANT: GANDSON, JEREMY
; APPLICANT: TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847B
; FILING DATE: 26-NO. 6582703-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 105:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1481 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 105:
US-08-979-847B-105

```

```

Query Match 1.0%; Score 73; DB 3; Length 1481;
Best Local Similarity 99.2%; Pred. No. 5.2e-26;
Matches 123; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6187 CCTGAACAATGGAAACAACCTTCAGCACAGAAATAAACACACATTCGCTTTTAGTAGACCT 6246
Db 607 CCTGAACAATGGAAACAACCTTCAGCACAGAAATAAACACACATTCGCTTTTAGTAGACCT 666

QY 6247 CTTGTTTCCAATSTGGAATAACCATACCTCAACCTCACCCTGTGTAATAATTAGCAAT 6306
Db 667 CTTGTTTCCAATCTGGAATAAACCATACCTCAACCTCACCCTGTGTAATAATTAGCAAT 726

QY 6307 ACTA 6310
Db 727 ACTA 730

```

```

RESULT 18
US-09-949-016-13002/c
; Sequence 13002, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307

```



; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13002  
; LENGTH: 168394  
; TYPE: DNA  
; ORGANISM: Human  
; NAME/KEY: misc feature  
; LOCATION: (1)...(168394)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-13002

Query Match 0.8%; Score 64; DB 3; Length 168394;  
Best Local Similarity 100.0%; Pred. No. 2e-21;  
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1999 TGAGTCAGAGCCACTAACCCAGATGATCCAGCAGCAGGACTGAGGGTGCTGGGCAAGC 2058  
Db 28029 TGAGTCAGAGCCACTAACCCAGATGATCCAGCAGCAGGACTGAGGGTGCTGGGCAAGC 27970  
QY 2059 GCCA 2062  
Db 27969 GCCA 27966

## RESULT 19

US-09-719-554-5/c  
; Sequence 5, Application US/09719554  
; Patent No. 6919438

; GENERAL INFORMATION:  
; APPLICANT: ALLIEL, Patrick  
; APPLICANT: PERIN, Jean-Pierre  
; APPLICANT: RIEGER, Francois  
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU  
; FILE REFERENCE: 200936US0PCT  
; CURRENT APPLICATION NUMBER: US/09/719,554  
; CURRENT FILING DATE: 2001-12-26  
; PRIOR APPLICATION NUMBER: PCT/FR99/01513  
; PRIOR FILING DATE: 1999-06-25  
; NUMBER OF SEQ ID NOS: 122  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 1799  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-719-554-5

Query Match 0.8%; Score 63; DB 3; Length 1799;  
Best Local Similarity 100.0%; Pred. No. 6e-21;  
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 6405 AGCCTATCGTTGTTGAATGGCTCTTCAGAACTCTATGTCCTCTCATTTAGTGCC 6464  
Db 938 AGCCTATCGTTGTTGAATGGCTCTTCAGAACTCTATGTCCTCTCATTTAGTGCC 879  
QY 6465 CCC 6467  
Db 878 CCC 876

## RESULT 20

US-09-949-016-15858/c  
; Sequence 15858, Application US/09949016  
; Patent No. 6812339

; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15858  
; LENGTH: 145320  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-15858

Query Match 0.8%; Score 63; DB 3; Length 145320;  
Best Local Similarity 100.0%; Pred. No. 6.5e-21;  
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 6405 AGCCTATCGTTGTTGAATGGCTCTTCAGAACTCTATGTCCTCTCATTTAGTGCC 6464  
Db 93202 AGCCTATCGTTGTTGAATGGCTCTTCAGAACTCTATGTCCTCTCATTTAGTGCC 93143  
QY 6465 CCC 6467  
Db 93142 CCC 93140

## RESULT 21

US-08-979-847B-112  
; Sequence 112, Application US/08979847B  
; Patent No. 6582703

; GENERAL INFORMATION:  
; APPLICANT: PERRON, HERVE  
; BESEME, FREDERIC  
; BEDIN, FREDERIC  
; PARANHOS-BACCALA, GLAUCIA  
; KOMURIAN-PRADEL, FLORENCE  
; JOLIVET-REYNAUD, COLETTE  
; MANDRAND, BERNARD  
; GARSON, JEREMY  
; TUBE, PHILIP

TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS  
ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHY  
THERAPEUTIC PURPOSES

NUMBER OF SEQUENCES: 210  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OLIFF & BERRIDGE, PLC  
STREET: P.O. BOX 19928  
CITY: ALEXANDRIA  
STATE: VA  
COUNTRY: USA  
ZIP: 22320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION NUMBER: US/08/979,847B  
FILING DATE: 26-No. 6582703-1997  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: BERRIDGE, WILLIAM P.  
REGISTRATION NUMBER: 30,024  
REFERENCE/DOCKET NUMBER: WPB 39046A  
TELECOMMUNICATION INFORMATION:

```
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 758 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 112:
US-08-979-847B-112

Query Match          0.8%; Score 58; DB 3; Length 758;
Best Local Similarity 100.0%; Pred. No. 2e-18; Indels 0; Gaps 0;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5199 GCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAACAT 5256
Db 645 GCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAACAT 702

RESULT 22
US-09-949-016-17417/c
; Sequence 17417, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17417
; LENGTH: 7772
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(7772)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17417

Query Match          0.7%; Score 56; DB 3; Length 7772;
Best Local Similarity 100.0%; Pred. No. 2.2e-17; Indels 0; Gaps 0;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5205 TTAGTTCGAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAACATTACA 5260
Db 16566 TTAGTTCGAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAACATTACA 16511

RESULT 23
US-09-949-016-12249/c
; Sequence 12249, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
```

```
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12249
; LENGTH: 7797
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(7797)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12249

Query Match          0.7%; Score 56; DB 3; Length 7797;
Best Local Similarity 100.0%; Pred. No. 2.2e-17; Indels 0; Gaps 0;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5205 TTAGTTCGAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAACATTACA 5260
Db 16570 TTAGTTCGAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAACATTACA 16515

RESULT 24
US-10-104-047-1148
; Sequence 1148, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: HL-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1148
; LENGTH: 2349
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1148

Query Match          0.7%; Score 55; DB 3; Length 2349;
Best Local Similarity 100.0%; Pred. No. 6.7e-17; Indels 0; Gaps 0;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7232 GGTTCCTCTGTGAGATGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGC 7286
Db 1626 GGTTCCTCTGTGAGATGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGC 1680

RESULT 25
US-09-949-016-13590/c
; Sequence 13590, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13590
```

```
;
; LENGTH: 149971
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(149971)
; OTHER INFORMATION: n = A,T,C or G
US-94-949-016-13590
```

Query Match 0.7%; Score 55; DB 3; Length 149971;  
Best Local Similarity 100.0%; Pred. No. 7.3e-17;  
Matches 55; Conservative 0; Mismatches 0; Indels 0

[illegible]

RESULT 26  
US-08-979-847B-117  
; Sequence 117, Application US/08979847B  
; Patent No. 6582703  
; GENERAL INFORMATION:  
; APPLICANT: PERRON, HERVE  
; BESEME, FREDERIC  
; BEDIN, FREDERIC  
; PARANHOS-BACCALA, GL  
; KOMURIAN-PRADEL, FLO  
; JOLIVET-REYNAUD, COL  
; MANDRAND, BERNARD  
; GARSON, JEREMY  
; TUKE, PHILIP

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

NUMBER OF SEQUENCES: 210  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OLIFF & BERRIDGE, PLC  
STREET: P.O. BOX 19928  
CITY: ALEXANDRIA  
STATE: VA  
COUNTRY: USA  
ZIP: 22320

```

;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
```

```
; ;  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/979,847B  
; FILING DATE: 26-No. 6582703-1997  
; CLASSIFICATION: <Unknown>  
;
```

ATTORNEY/AGENT INFORMATION:  
NAME: BERRIDGE, WILLIAM P.  
REGISTRATION NUMBER: 30,024  
REFERENCE/DOCKET NUMBER: WPB 39046A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-836-6400  
TELEFAX: 703-836-2787

```

? INFORMATION FOR SEQ ID NO: 117:
?     SEQUENCE CHARACTERISTICS:
?         LENGTH: 1511 base pairs
?         TYPE: nucleic acid
?         STRANDEDNESS: single
?         TOPOLOGY: linear

```

```

; MOLECULE TYPE: CDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 117:
US-08-979-847B-117

```

```
Query Match      0.7%; Score 54; DB 3; Length 1511;
Best Local Similarity 100.0%; Pred. No. 2.1e-16;
Matches 54; Conservative 0; Mismatches 0; Indels
```

Qy	1615	TAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAGCTTGCC
Db	1003	TAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAGCTTGCC

```

RESULT 27
US-09-403-343B-32
; Sequence 32, Application US/09403343B
; Patent No. 6555091
; GENERAL INFORMATION:
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: PERRON, HERVE
; APPLICANT: MANDRAND, BERNARD
; TITLE OF INVENTION: POLYPEPTIDE CAPABLE OF
; TITLE OF INVENTION: PATIENTS SUFFERING FROM
; FILE REFERENCE: 104574
; CURRENT APPLICATION NUMBER: US/09/403,343B
; CURRENT FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: FR/97/05679
; PRIOR FILING DATE: 1997-04-29
; PRIOR APPLICATION NUMBER: FR/97/16870
; PRIOR FILING DATE: 1997-12-31
; PRIOR APPLICATION NUMBER: PCT/FR98/00870
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 286
; TYPE: DNA
; ORGANISM: MSRV
US-09-403-343B-32

```

```
Query Match          .      0.7%; Score 53; DB 3; Length 286;
Best Local Similarity 100.0%; Pred. No. 6.6e-16;
Matches 53: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

**Qy** 1513 GGAAGAAATCCCTACTGCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGC 1565  
|||||  
**Dd** 12 GGAAGAAATCCCTACTGCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGC 64

RESULT 28  
US-08-691-563C-88  
; Sequence 88, Application US/08691563C  
; Patent No. 6001987  
; GENERAL INFORMATION:

/ APPLICANT: Herve PERRON  
 / APPLICANT: Frederic BESEME  
 / APPLICANT: Frederic BEDIN  
 / APPLICANT: GLAUCIA PARANHOS-BACCALA  
 / APPLICANT: FLORENCE KOMURIAN-PRADEL  
 / APPLICANT: Colette JOLIVET  
 / APPLICANT: Bernard MANDRAND  
 / TITLE OF INVENTION: VIRAL MATERIAL /  
 / TITLE OF INVENTION: ASSOCIATED WITH /  
 / TITLE OF INVENTION: THERAPEUTIC PUR /  
 / NUMBER OF SEQUENCES: 92  
 / CORRESPONDENCE ADDRESS:

CONFERENCE ADDRESS:  
ADDRESSEE: Oliff & Berridge  
STREET: 700 South Washington Street, Suite 300  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: U.S.A.

ZIP: 22314  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/691,563C  
FILING DATE: 02-AUG-1996  
ATTORNEY/AGENT INFORMATION:

```
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 38588
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 88:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 693 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-691-563C-88

Query Match      0.7%; Score 53; DB 3; Length 693;
Best Local Similarity 100.0%; Pred. No. 6.7e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1513 GGAAGAAATCCTACTGCGCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGC 1565
      |||||||
Db 577 GGAAGAAATCCTACTGCGCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGC 629

RESULT 29
US-09-374-766-88
; Sequence 88, Application US/09374766
; Patent No. 6579526
; GENERAL INFORMATION:
; APPLICANT: Herve PERRON
; APPLICANT: Frederic BESEME
; APPLICANT: Frederic BEDIN
; APPLICANT: Glauclia PARANHOS-BACCALA
; APPLICANT: Florence KOMURIAN-PRADEL
; APPLICANT: Colette JOLIVET
; APPLICANT: Bernard MANDRAND
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 92
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oliff & Berridge
; STREET: 700 South Washington Street, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: US/09/374,766
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/691,563
; FILING DATE: 02-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 38588
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 88:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 693 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-374-766-88
```

```
Query Match      0.7%; Score 53; DB 3; Length 693;
Best Local Similarity 100.0%; Pred. No. 6.7e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1513 GGAAGAAATCCTACTGCGCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGC 1565
      |||||||
Db 577 GGAAGAAATCCTACTGCGCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGC 629

RESULT 30
US-08-979-847B-82
; Sequence 82, Application US/08979847B
; Patent No. 6582703
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; APPLICANT: BESEME, FREDERIC
; APPLICANT: BEDIN, FREDERIC
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: KOMURIAN-PRADEL, FLORENCE
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: MANDRAND, BERNARD
; APPLICANT: GARSON, JEREMY
; APPLICANT: TUKES, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847B
; FILING DATE: 26-No. 6582703-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 82:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 693 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 82:
; US-08-979-847B-82

Query Match      0.7%; Score 53; DB 3; Length 693;
Best Local Similarity 100.0%; Pred. No. 6.7e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1513 GGAAGAAATCCTACTGCGCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGC 1565
      |||||||
Db 577 GGAAGAAATCCTACTGCGCTTTCTGGAGAGACTAAGGGAGGCATTGAGGAAGC 629

Search completed: August 2, 2006, 13:30:03
Job time : 1256 secs
```

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2006, 09:10:55 ; Search time 21273 Seconds  
(without alignments)  
19930.446 Million cell updates/sec

Title: US-10-717-580-11  
Perfect score: 7582  
Sequence: 1 caacatcggtataaac.....tattaaattgcactgtg 7582

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0  
Searched: 48236798 seqs, 27959665780 residues

Word size : 6  
Total number of hits satisfying chosen parameters: 96469239

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 90 summaries

Database : EST:  
1: gb\_est1:  
2: gb\_est3:  
3: gb\_est4:  
4: gb\_est5:  
5: gb\_est6:  
6: gb\_hic:  
7: gb\_est2:  
8: gb\_est7:  
9: gb\_est8:  
10: gb\_est9:  
11: gb\_gss1:  
12: gb\_gss2:  
13: gb\_gss3:  
14: gb\_gss4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	731	9.6	2500	6	CR622175 full-leng
2	731	9.6	2716	6	CR613169 full-leng
3	731	9.6	2748	6	CR605851 full-leng
4	731	9.6	2748	6	CR625046 full-leng
5	731	9.6	2749	6	CR617248 full-leng
6	525	6.9	935	4	CR617248 full-leng
7	446	5.9	828	4	CR617248 full-leng
8	442	5.8	903	4	CR617248 full-leng
9	437	5.8	883	4	CR617248 full-leng
10	431	5.7	872	4	CR617248 full-leng
11	411	5.4	931	4	CR617248 full-leng
12	403	5.3	792	9	CR617248 full-leng
13	379	5.0	564	9	CR617248 full-leng
14	377	5.0	924	4	CR617248 full-leng
15	374	4.9	538	9	CR617248 full-leng
16	373	4.9	721	1	CR617248 full-leng
17	373	4.9	761	9	CR617248 full-leng
18	372	4.9	540	9	CR617248 full-leng
19	370	4.9	819	9	CR617248 full-leng

20	369	4.9	921	4	BX408734
21	367	4.8	571	9	DA846799
22	363	4.8	541	9	DA835080
23	360	4.7	846	4	BX368078
24	358	4.7	547	9	DA845887
25	357	4.7	571	9	DA846543
26	356	4.7	995	4	BX389657
27	353	4.7	658	4	BX430050
28	344	4.5	955	4	BX347952
29	338	4.5	554	8	CR735592
30	336	4.4	924	4	BX409328
31	324	4.3	586	7	BE019603
32	309	4.1	702	1	AU138097
33	309	4.1	844	4	BX347111
34	302	4.0	814	4	BX430055
35	300	4.0	836	4	BX388766
36	298	3.9	790	2	EG572445
37	295	3.9	503	2	BG000981
38	295	3.9	578	9	DA845799
39	292	3.9	877	4	BX347619
40	290	3.8	769	9	DA833923
41	279	3.7	717	9	DA830985
42	273	3.6	376	2	BI056595
43	271	3.6	373	2	BI052559
44	271	3.6	373	2	BI052567
45	269	3.5	547	9	DA838367
46	266	3.5	379	2	BI052566
47	253	3.3	698	9	DA834382
48	251	3.3	572	9	DA851356
49	250	3.3	572	9	DA851363
50	249	3.3	427	1	AA776439
51	247	3.3	366	2	BF993782
52	247	3.3	770	9	DR004257
53	246	3.2	373	2	BI052561
54	246	3.2	562	9	DA849456
55	239	3.2	775	9	DA834377
56	224	3.0	414	10	HI2439
57	224	3.0	459	10	R68740
58	221	2.9	374	2	BI056593
59	219	2.9	362	2	BI049622
60	216	2.8	373	2	BI052571
61	216	2.8	373	2	BI056589
62	213	2.8	539	9	DB068120
63	213	2.8	543	9	DB062480
64	211	2.8	552	2	DA838102
65	210	2.8	373	2	BI052570
66	204	2.7	701	9	DA834293
67	200	2.6	313	2	BI056596
68	199	2.6	905	4	BX429316
69	197	2.6	564	4	BX388917
70	196	2.6	298	2	BF992540
71	190	2.5	562	9	DA835505
72	190	2.5	578	9	DA848644
73	190	2.5	872	4	DA8459154
74	190	2.5	936	4	BX337770
75	188	2.5	944	4	BX368759
76	181	2.4	317	2	BI056590
77	180	2.4	379	9	D78692
78	177	2.3	758	4	DX357208
79	177	2.3	1071	4	BX365066
80	176	2.3	724	9	DA834354
81	174	2.3	382	2	BI052569
82	173	2.3	484	4	BX409035
83	171	2.3	567	9	DA861649
84	171	2.3	797	4	CB990962
85	171	2.3	803	4	CB992422
86	171	2.3	823	4	CB993677
87	171	2.3	826	4	CB990802
88	171	2.3	837	4	CB996320
89	171	2.3	852	4	CB994833
90	171	2.3	1160	6	BC030968

20	369	4.9	921	4	BX408734
21	367	4.8	571	9	DA846799
22	363	4.8	541	9	DA835080
23	360	4.7	846	4	BX368078
24	358	4.7	547	9	DA845887
25	357	4.7	571	9	DA846543
26	356	4.7	995	4	BX389657
27	353	4.7	658	4	BX430050
28	344	4.5	955	4	BX347952
29	338	4.5	554	8	CR735592
30	336	4.4	924	4	BX409328
31	324	4.3	586	7	BE019603
32	309	4.1	702	1	AU138097
33	309	4.1	844	4	BX347111
34	302	4.0	814	4	BX430055
35	300	4.0	836	4	BX388766
36	298	3.9	790	2	EG572445
37	295	3.9	503	2	BG000981
38	295	3.9	578	9	DA845799
39	292	3.9	877	4	BX347619
40	290	3.8	769	9	DA833923
41	279	3.7	717	9	DA830985
42	273	3.6	376	2	BI056595
43	271	3.6	373	2	BI052559
44	271	3.6	373	2	BI052567
45	269	3.5	547	9	DA838367
46	266	3.5	379	2	BI052566
47	253	3.3	698	9	DA834382
48	251	3.3	572	9	DA851356
49	250	3.3	572	9	DA851363
50	249	3.3	427	1	AA776439
51	247	3.3	366	2	BF993782
52	247	3.3	770	9	DR004257
53	246	3.2	373	2	BI052561
54	246	3.2	562	9	DA849456
55	239	3.2	775	9	DA834377
56	224	3.0	414	10	HI2439
57	224	3.0	459	10	R68740
58	221	2.9	374	2	BI056593
59	219	2.9	362	2	BI049622
60	216	2.8	373	2	BI052571
61	216	2.8	373	2	BI056589
62	213	2.8	539	9	DB068120
63	213	2.8	543	9	DB062480
64	211	2.8	552	2	DA838102
65	210	2.8	373	2	BI052570
66	204	2.7	701	9	DA834293
67	200	2.6	313	2	BI056596
68	199	2.6	905	4	BX429316
69	197	2.6	564	4	BX388917
70	196	2.6	298	2	BF992540
71	190	2.5	562	9	DA835505
72	190	2.5	578	9	DA848644
73	190	2.5	872	4	DA8459154
74	190	2.5	936	4	BX337770
75	188	2.5	944	4	BX368759
76	181	2.4	317	2	BI056590
77	180	2.4	379	9	D78692
78	177	2.3	758	4	DX357208
79	177	2.3	1071	4	BX365066
80	176	2.3	724	9	DA834354
81	174	2.3	382	2	BI052569
82	173	2.3	484	4	BX409035
83	171	2.3	567	9	DA861649
84	171	2.3	797	4	CB990962
85	171	2.3	803	4	CB992422
86	171	2.3	823	4	CB993677
87	171	2.3	826	4	CB990802
88	171	2.3	837	4	CB996320
89	171	2.3	852	4	CB994833
90	171	2.3	1160	6	BC030968

## ALIGNMENTS

```
RESULT 1
CR622175      2500 bp      mRNA      linear      HTC 21-JUL-2004
LOCUS         full-length cDNA clone CS0DI051YM13 of Placenta Cot 25-normalized
DEFINITION    of Homo sapiens (human) .
ACCESSION    CR622175
VERSION      CR622175.1 GI:50502982
KEYWORDS     HTC; CNSLT cDNA.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Hominiidae; Homo.
              1 (bases 1 to 2500)
REFERENCE    Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
AUTHORS     Full-length cDNA libraries and normalization
TITLE       Unpublished
JOURNAL     Contact : Feng Liang Email : fliang@lifetech.com URL :
REMARK      http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
              Faraday Avenue
              2 (bases 1 to 2500)
REFERENCE    Genoscope.
AUTHORS     Direct Submission
TITLE       Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL     BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
              - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
              end enriched, double-strand cDNA was digested with Not I and cloned
              into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
              was normalized. Library was constructed by Life Technologies, a
              division of Invitrogen.
FEATURES     Location/Qualifiers
              source          1..2500
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="CS0DI051YM13"
                        /tissue type="Placenta Cot 25-normalized"
                        /plasmid="pCMVSPORT_6"
ORIGIN
Query Match          9.6%; Score 731; DB 6; Length 2500;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 5685 AGAGTTTCTATGGAGATCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 5744
DB      |||||||
QY 634 AGAGTTTCTATGGAGATCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 693
DB      |||||||
QY 5745 TTSTAGGGAACCCACCTTCACTGCCACACCCCATATGCCCGCAACTGCTATCACTC 5804
DB      |||||||
QY 694 TTCTAAGGGAACCCACCTTCACTGCCACACCCCATATGCCCGCAACTGCTATCACTC 753
DB      |||||||
QY 5805 TGCCACTTTTGATCGATGCAAACTCACTATTATTGGACAGGAAATGATTAATCTTAG 5864
DB      |||||||
QY 754 TGCCACTTTTGATCGATGCAAACTCACTATTATTGGACAGGAAATGATTAATCTTAG 813
DB      |||||||
QY 5865 TTGTCTCGGAGGACTTGGAGTCACTGCTGTGGACTTACTTCAACCACCTGGTATGTC 5924
DB      |||||||
QY 814 TTGTCTCGGAGGACTTGGAGTCACTGCTGTGGACTTACTTCAACCACCTGGTATGTC 873
DB      |||||||
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAGAAACATGTAAAGAAAGTAACTCCCA 5984
DB      |||||||
QY 874 TGTATGGGGTGGAGTTCAAGATCAGGCAAGAGAGAAACATGTAAAGAAAGTAACTCCCA 933
DB      |||||||
QY 5985 ACTACCSGGGTACATGGCACTCTTAGGCCCTCAAGAGACTAGATCTCTCAAAACTACA 6044
DB      |||||||
QY 934 ACTACCSGGGTACATGGCACTCTTAGGCCCTCAAGAGACTAGATCTCTCAAAACTACA 993
DB      |||||||
QY 6045 TGAACCCCTCCGTACCCTACTCGCTGGTAAAGCCTATTATTAATACCACCTCACTGGGCT 6104
DB      |||||||
```

```
Db 994 TGAACCCCTCCGTACCCTACTCGCTGGTAAAGCCTATTATTAATACCACCTCACTGGGCT 1053
QY 6105 CCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTTGGATGCTCCCTCCCTGGAATTCAR 6164
DB 1054 CCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTTGGATGCTCCCTCCCTGGAATTCAG 1113
QY 6165 GCCATATGTTTCAATCCCTGTACCTGAACTGAACTGAACTGAACTGAACTGAACTGAACT 6224
DB 1114 GCCATATGTTTCAATCCCTGTACCTGAACTGAACTGAACTGAACTGAACTGAACTGAACT 1173
QY 6225 CACTTCCGTTTGTAGTAGGACCTCTTGTTCCTCAATSTGGAATAAACCCATACCTCAAACT 6284
DB 1174 CACTTCCGTTTGTAGTAGGACCTCTTGTTCCTCAATSTGGAATAAACCCATACCTCAAACT 1233
QY 6285 CACTGTGTAAATTTAGCAATATATACATACAACTCCCAATGATGATGAGTGGT 6344
DB 1234 CACTGTGTAAATTTAGCAATATATACATACAACTCCCAATGATGATGAGTGGT 1293
QY 6345 AACTCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTC 6404
DB 1294 AACTCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTC 1353
QY 6405 AGCTTATCTGTTTGAATGGCTCTTCAAGAACTATATGTCTCTCTCATTTCTAGTGC 6464
DB 1354 AGCTTATCTGTTTGAATGGCTCTTCAAGAACTATATGTCTCTCTCATTTCTAGTGC 1413
QY 6465 CCCATGATCCATCTACACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACT 6524
DB 1414 CCCTATGACCATCTACACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACT 1473
QY 6525 CAAAGAGTACCCATCTTCTCTTGTATAGGACAGAGTGTAGTGCAGTACAGTAC 6584
DB 1474 CAAAGAGTACCCATCTTCTCTTGTATAGGACAGAGTGTAGTGCAGTACAGTAC 1533
QY 6585 TGGCATTTGGCGGTATCAAACTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAA 6644
DB 1534 TGGCATTTGGCGGTATCAAACTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAA 1593
QY 6645 TGGGACATGAAAGCGGTGCGGACTCCCTGGTCACTCTGGTCACTTGCAGATCACTTACCT 6704
DB 1594 TGGGACATGAAAGCGGTGCGGACTCCCTGGTCACTCTGGTCACTTGCAGATCACTTACCT 1653
QY 6705 AGCAGCAGTGTCTTC 6721
DB 1654 AGCAGCAGTGTCTTC 1670

RESULT 2
CR613169
LOCUS         full-length cDNA clone CS0DE013YI20 of Placenta of Homo sapiens
DEFINITION    (human) .
ACCESSION    CR613169
VERSION      CR613169.1 GI:50493976
KEYWORDS     HTC; CNSLT cDNA.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Hominiidae; Homo.
              1 (bases 1 to 2716)
REFERENCE    Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
AUTHORS     Full-length cDNA libraries and normalization
TITLE       Unpublished
JOURNAL     Contact : Feng Liang Email : fliang@lifetech.com URL :
REMARK      http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
              Faraday Avenue
              2 (bases 1 to 2716)
REFERENCE    Genoscope.
AUTHORS     Direct Submission
TITLE       Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL     BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
```







```
QY 6225 CACTTCCGTTTGTAGTACGACTCTTGTTCCTCAATSTGGAATACCATACCTCCTCAACCT 6284
Db 1427 CACTTCCGTTTGTAGTACGACTCTTGTTCCTCAATSTGGAATACCATACCTCCTCAACCT 1486
QY 6285 CACTTGTGTAAATTTAGCAATACATACACAAACCAACTCCCAATGCATCAGTGGGT 6344
Db 1487 CACTTGTGTAAATTTAGCAATACATACACAAACCAACTCCCAATGCATCAGTGGGT 1546
QY 6345 AACTTCTCCACACAAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGTGCTACCTC 6404
Db 1547 AACTTCTCCACACAAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGTGCTACCTC 1606
QY 6405 AGCTATCGTTTGTGAATGGCTCTTCAGAAATCTATGTGCTCTCTCATCTTCTAGTGCC 6464
Db 1607 AGCTATCGTTTGTGAATGGCTCTTCAGAAATCTATGTGCTCTCTCATCTTCTAGTGCC 1666
QY 6465 CCVATGRCATCTACACTGAACAAAGATTTATACAGTTATGTCTATCTAAGCCCGCAA 6524
Db 1667 CCCTATGACCATCTACACTGAACAAAGATTTATACAGTTATGTCTATCTAAGCCCGCAA 1726
QY 6525 CAAGAAGTACCCATCTCTCTTCTTGTATAGAGCAGGAGTGTAGGTGCATAGGTAC 6584
Db 1727 CAAGAAGTACCCATCTCTCTTCTTGTATAGAGCAGGAGTGTAGGTGCATAGGTAC 1786
QY 6585 TGGCATTGGCGGTATCAACACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 6644
Db 1787 TGGCATTGGCGGTATCAACACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 1846
QY 6645 TGGGACATGGAACGGGTGCGGACTCCCTGGTCACTTCAAGAGTCAAACTTAACTCCCT 6704
Db 1847 TGGGACATGGAACGGGTGCGGACTCCCTGGTCACTTCAAGAGTCAAACTTAACTCCCT 1906
QY 6705 AGCAGCATGAGTCTTCTC 6721
Db 1907 AGCAGCATGAGTCTTCTC 1923

RESULT 5
LOCUS CR617248
DEFINITION full-length cDNA clone CS0D1022YJ18 of Placenta HTC 21-JUL-2004
of Homo sapiens (human).
ACCESSION CR617248
VERSION 1
KEYWORDS full-length cDNA libraries and normalization
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 2749)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 2749)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail: segref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..2749
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
```

```
ORIGIN
Query Match 9.6%; Score 731; DB 6; Length 2749;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 5685 AGAGTTTCTATGAGGAATGCAGCGTCCCGGAATATTTGATGCCCAATCGTATAGGAGTCT 5744
Db 881 AGAGTTTCTATGAGGAATGCAGCGTCCCGGAATATTTGATGCCCAATCGTATAGGAGTCT 940
QY 5745 TTSTAGGGAACCCCACTTCACTGCTGCCACACCCATATGCTCCCGCAACTGCTATCATCTC 5804
Db 941 TTCTAAGGGAACCCCACTTCACTGCTGCCACACCCATATGCTCCCGCAACTGCTATCATCTC 1000
QY 5805 TGCACCTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5864
Db 1001 TGCACCTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1060
QY 5865 TTGTCTCTGGAGGACTTGGAGTCACTGCTGCTGTTGGACTTACTTCAACCAAACTGGTATGTC 5924
Db 1061 TTGTCTCTGGAGGACTTGGAGTCACTGCTGCTGTTGGACTTACTTCAACCAAACTGGTATGTC 1120
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAAGAAATTAATCTCCCA 5984
Db 1121 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAAGAAATTAATCTCCCA 1180
QY 5985 ACTCACSGGGTACATGCGACCTCTAGCCCTCAAGAGGACTAGATCTCTCAAAACTACA 6044
Db 1181 ACTCACSGGGTACATGCGACCTCTAGCCCTCAAGAGGACTAGATCTCTCAAAACTACA 1240
QY 6045 TGAACCCCTCCGTACCCATACCTCGCTGGTAAAGCTTAAATACCAACCTCTCTGGGT 6104
Db 1241 TGAACCCCTCCGTACCCATACCTCGCTGGTAAAGCTTAAATACCAACCTCTCTGGGT 1300
QY 6105 CCATGAGGTCTGGGCCCAAAACCCCTACTAATCTTGGATATGCTCCCTGAACTTTCA 6164
Db 1301 CCATGAGGTCTGGGCCCAAAACCCCTACTAATCTTGGATATGCTCCCTGAACTTTCA 1360
QY 6165 GCCATATGTTTCAATCCCTGTACTGCAACAATGGAACTTCCAGACAGAAATAAACAC 6224
Db 1361 GCCATATGTTTCAATCCCTGTACTGCAACAATGGAACTTCCAGACAGAAATAAACAC 1420
QY 6225 CACTTCCGTTTGTAGTAGGACCTCTGCTTTTCCAAATSTGGAATTAACCCATACCTCAACCT 6284
Db 1421 CACTTCCGTTTGTAGTAGGACCTCTGCTTTTCCAAATSTGGAATTAACCCATACCTCAACCT 1480
QY 6285 CACTTGTGTAAATTTAGCAATACATACACAAACCAACTCCCAATGCATCAGGTGGGT 6344
Db 1481 CACTTGTGTAAATTTAGCAATACATACACAAACCAACTCCCAATGCATCAGGTGGGT 1540
QY 6345 AACTCTCCCAACAAATAGTCTGCCCTACCTCAGGAATATTTTGTCTGTGTGCTACCTC 6404
Db 1541 AACTCTCCCAACAAATAGTCTGCCCTACCTCAGGAATATTTTGTCTGTGTGCTACCTC 1600
QY 6405 AGCTATCGTTTGTGAATGGCTCTTCAGAAATCTATGTGCTCTCTCATCTTCTAGTGCC 6464
Db 1601 AGCTATCGTTTGTGAATGGCTCTTCAGAAATCTATGTGCTCTCTCATCTTCTAGTGCC 1660
QY 6465 CCCTATGRCATCTACACTGAACAAAGATTTATACAGTTATGTCTATCTAAGCCCGCAA 6524
Db 1661 CCCTATGRCATCTACACTGAACAAAGATTTATACAGTTATGTCTATCTAAGCCCGCAA 1720
QY 6525 CAAAAGAGTACCCATCTTCTCTTTTGTATAGAGCAGGAGTGTAGGTGCATAGGTAC 6584
Db 1721 CAAAAGAGTACCCATCTTCTCTTTTGTATAGAGCAGGAGTGTAGGTGCATAGGTAC 1780
QY 6585 TGGCATTTGGGGTATCAGAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 6644
Db 1781 TGGCATTTGGGGTATCAGAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 1840
```

```
/clone="CS0D1022YJ18"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"
```

```
QY 6645 TGGGACATGGAACGGGTGCGGACTCCTCGTGTACCTTGGTCACTTGAAGATCAACTTAACCTCCT 6704
|||||
Db 1841 TGGGACATGGAACGGGTGCGGACTCCTCGTGTACCTTGGTCACTTGAAGATCAACTTAACCTCCT 1900

QY 6705 AGCAGCAGTAGTCCTTC 6721
|||||
Db 1901 AGCAGCAGTAGTCCTTC 1917

RESULT 6
BX391741/c 935 bp mRNA linear EST 29-APR-2004
DEFINITION BX391741 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS01051YM13 3-PRIME, mRNA sequence.
ACCESSION BX391741
VERSION BX391741.1 GI:30619529
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 935)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BA1022ZA10_CS02089_1&c=4215.r
```

## FEATURES

```
source
Location/Qualifiers
1..935
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS01051YM13"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 6.9%; Score 525; DB 4; Length 935;
Best Local Similarity 99.5%; Pred. No. 8.3e-282; Indels 0; Gaps 0;
Matches 725; Conservative 0; Mismatches 4;
```

```
QY 5993 GGGTACATGGCACCTTAGCCCTACAAAGGACTAGACTCTCAAAACTACATGAACCC 6052
|||||
Db 746 GGGTACATGGCACCTTAGCCCTACAAAGGACTAGACTCTCAAAACTACATGAACCC 687

QY 6053 TCCGTACCATACTCGCTGGTAAGCCTATTTTAATACCACTCTACTGGGCTCCATGAGG 6112
|||||
Db 686 TCCGTACCATACTCGCTGGTAAGCCTATTTTAATACCACTCTACTGGGCTCCATGAGG 627

QY 6113 TCTCGSCCAAAACCTACTACTGTGGATATGCTCCCTGCACTTCARGCCATATG 6172
|||||
Db 626 TCTCGSCCAAAACCTACTACTGTGGATATGCTCCCTGCACTTCARGCCATATG 567

QY 6173 TTTCAATCCCTGTACTGGAACCAATGGAACAACCTTCAGCACAGAAATAAACACCTTCG 6232
|||||
Db 566 TTTCAATCCCTGTACTGGAACCAATGGAACAACCTTCAGCACAGAAATAAACACCTTCG 507
```

```
QY 6233 TTTTAGTAGGACCTCTTGTCTTCCAAATGTGAAATAAACCCATCACTCAAACTCACCTGTG 6292
|||||
Db 506 TTTTAGTAGGACCTCTTGTCTTCCAAATGTGAAATAAACCCATCACTCAAACTCACCTGTG 447

QY 6293 TAAATTTAGCAATATCTACATACACCAACCAATGCCAATGCATCAGGTGGGTAACCTC 6352
|||||
Db 446 TAAATTTAGCAATATCTACATACACCAACCAATGCCAATGCATCAGGTGGGTAACCTC 387

QY 6353 CCACACAAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATC 6412
|||||
Db 386 CCACACAAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATC 327

QY 6413 GTTGTTCGAATGGCTCTTCAGAACTATATGCTCTCTCATTTCTTAGTGCCCTCATGR 6472
|||||
Db 326 GTTGTTCGAATGGCTCTTCAGAACTATATGCTCTCTCATTTCTTAGTGCCCTCATGA 267

QY 6473 CCATCTACACTGAACAAGATTATACAGTTATGTGATATCTAAGCCCGCAACAAAGAG 6532
|||||
Db 266 CCATCTACACTGAACAAGATTATACAGTTATGTGATATCTAAGCCCGCAACAAAGAG 207

QY 6533 TACCATTCTTCTTTTGTATAGAGCAGAGTGCTAGGTGCATAGGTACTGCATTTG 6592
|||||
Db 206 TACCATTCTTCTTTTGTATAGAGCAGAGTGCTAGGTGCATAGGTACTGCATTTG 147

QY 6593 GCGGTATCACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACA 6652
|||||
Db 146 GCGGTATCACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACA 87

QY 6653 TGGAAAGCGGTGCGGACTCCCTGGTTCACCTTTCGAAGATCAACTTAACCTCCCTAGCAGCAG 6712
|||||
Db 86 TGGAAAGCGGTGCGGACTCCCTGGTTCACCTTTCGAAGATCAACTTAACCTCCCTAGCAGCAG 27

QY 6713 TAGTCTTTC 6721
|||||
Db 26 TAGTCTTTC 18
```

## RESULT 7

```
LOCUS BX367907
DEFINITION BX367907 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS01051YM13 5-PRIME, mRNA sequence.
ACCESSION BX367907
VERSION BX367907.1 GI:30459586
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 828)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0AT022ZG01_T02023_1&c=4215.r.
```

## FEATURES

```
source
Location/Qualifiers
1..828
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS01051YM13"
/tissue_type="PLACENTA COT 25-NORMALIZED"
```

```
/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
```

Query Match	5.9%;	Score 446;	DB 4;	Length 828;
Best Local Similarity	99.4%;	Pred. No. 2e-237;		
Matches 646;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;
QY	5748	TAAGGGAAACCCACCTTCACTGCGCCACACCCCATATGCCCCGCAACTGCTATCACTCTGC	5807	
Db	66	TAAGGGAAACCCACCTTCACTGCGCCACACCCCATATGCCCCGCAACTGCTATCACTCTGC	125	
QY	5808	CACCTCTTTGCATGTCATCAAACTACTCAATATTGGACAGGAAAAATGATTAATCCTAGCTTG	5867	
Db	126	CACCTCTTTGCATGTCATCAAACTACTCAATATTGGACAGGAAAAATGATTAATCCTAGCTTG	185	
QY	5868	TCTCGAGGACTTGGAGTCACCTGCTGTGTGGACTTACTTCACCCAAACTGGTATGTCGA	5927	
Db	186	TCTCGAGGACTTGGAGTCACCTGCTGTGTGGACTTACTTCACCCAAACTGGTATGTCGA	245	
QY	5928	TGGGGTGGAGTTCAGATCAGGCAAGAGAAAAACAINGTAAAAAGAGTAATTCCTCCAACT	5987	
Db	246	TGGGGTGGAGTTCAGATCAGGCAAGAGAAAAACAINGTAAAAAGAGTAATTCCTCCAACT	305	
QY	5988	CACCSGGTACATGGGCACCTCTAGCCCTACAAGGACTAGATCTCTCAAAACTACATGA	6047	
Db	306	CACCSGGTACATGGGCACCTCTAGCCCTACAAGGACTAGATCTCTCAAAACTACATGA	365	
QY	6048	AACCTCTCGGTACCCATACCTCGCTGGTAAAGCCTATTTAATACCAACCCTCACTGGGCTCCA	6107	
Db	366	AACCTCTCGGTACCCATACCTCGCTGGTAAAGCCTATTTAATACCAACCCTCACTGGGCTCCA	425	
QY	6108	TGAGGTCTCGGCCCAAAAACCTACTAACTGTGTGGATATGCCTCCCCCTGAACCTTCARGCC	6167	
Db	426	TGAGGTCTCGGCCCAAAAACCTACTAACTGTGTGGATATGCCTCCCCCTGAACCTTCARGCC	485	
QY	6168	ATATGTTTCAATCCCTGTACTTGAAACAATGGAGCAACTTCAGACAGAAATAAACACCCAC	6227	
Db	486	ATATGTTTCAATCCCTGTACTTGAAACAATGGAGCAACTTCAGACAGAAATAAACACCCAC	545	
QY	6228	TTCCGTTTTAGTAGGACCTCTGTGTTTCCAATSTGGAAATAACCCATACCTCAAAACCTCAC	6287	
Db	546	TTCCGTTTTAGTAGGACCTCTGTGTTTCCAATCTGGAAATAACCCATACCTCAAAACCTCAC	605	
QY	6288	CTGTGTAAATTTAGCAATACTACATACAACAACCAACTCCCAATGCATCAGGTGGGTAAAC	6347	
Db	606	CTGTGTAAATTTAGCAATACTACATACAACAACCAACTCCCAATGCATCAGGTGGGTAAAC	665	
QY	6348	TCTCCCAACAAATAGTCTGCCTACCTCAGGAATATTTTTGTCTGTG	6397	
Db	666	TCTCCCAACAAATAGTCTGCCTACCTCAGGAATATTTTTGTCTGTG	715	

RESULT 8	
BX408752	
LOCUS	903 bp mRNA linear EST 01-MAY-2004
DEFINITION	Homo sapiens PLACENTA Homo sapiens cDNA clone CSODE013Y120 5-PRIME, mRNA sequence.
ACCESSION	BX408752
VERSION	BX408752.2 GI:46931100
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 903)
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)

**COMMENT**

On May 13, 2003 this sequence version replaced gi:30648825.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE  
E mail: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NorI-oligo(dT) primer. Five prime  
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
was not normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.

This sequence belongs to sequence cluster 4215.  
For more information about this cluster, see  
<http://www.genoscope.cns.fr/cdna?e=CS0BAA008ZH03>

## FEATURES

```

Location/Qualifiers
1. .903
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS00E013Y1212"
/tissue_type="PLACENTA"
/clone_libs="Homo sapiens"
/notes="Vector: pCMVcat
with a NotI-oligo (d(GAATTC)
double-strand cDNA library
the Not I and EcoRV sites
Library was not normalized"

```

## ORIGIN

Query Match	5.8%;	Score 442;	DB 4;	Length 903;
Best Local Similarity	99.3%;	Pred. No. 3.5e-235;		
Matches 692;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;
QY	5685	AGAGTTTTCTATGAGAAATGCACGCTCCCGGAAATATTGATGCCCGCAATCTGTATAGGAGTCT	5744	
Db	151	AGAGTTTTCTATGAGAAATGCACGCTCCCGGAAATATTGATGCCCGCAATCTGTATAGGAGTCT	210	
QY	5745	TTSTAAGGGAACCCCAACCTTCACCTGCCACACCCCATATGCCCGCCCAACTGCTATCACCTC	5804	
Db	211	TTCTAAGGGAACCCCAACCTTCACCTGCCACACCCCATATGCCCGCCCAACTGCTATCACCTC	270	
QY	5805	TGCCACTTTTGCATGCAATGCAAAATACATTTATTGGACAGGAAAAATGATTAAATCCTTAG	5864	
Db	271	TGCCACTTTTGCATGCAATGCAAAATACATTTATTGGACAGGAAAAATGATTAAATCCTTAG	330	
QY	5865	TTGTCTCGGAGGACTTGGAGTCACTGTCTGTGTGGACTTACTTCACCCAAACTGGTATGTC	5924	
Db	331	TTGTCTCGGAGGACTTGGAGTCACTGTCTGTGTGGACTTACTTCACCCAAACTGGTATGTC	390	
QY	5925	TGATGGGGTGGAGTTCAAGATCAGCGCAAGAGAAAAACATGTAAAGAAAGTAAATCTCCCA	5984	
Db	391	TGATGGGGTGGAGTTCAAGATCAGCGCAAGAGAAAAACATGTAAAGAAAGTAAATCTCCCA	450	
QY	5985	ACTCACCSGGGTACATGCGACCTCTAGGCCCTACAAAGAGCTAGATCTCTCAAAACTACA	6044	
Db	451	ACTCACCSGGGTACATGCGACCTCTAGGCCCTACAAAGAGCTAGATCTCTCAAAACTACA	510	
QY	6045	TGAAACCCCTCGGTACCCATCTACGCTTGGTAAAGCCTATTTAATACCAACCCCTCACTGGGCT	6104	
Db	511	TGAAACCCCTCGGTACCCATCTACGCTTGGTAAAGCCTATTTAATACCAACCCCTCACTGGGCT	570	
QY	6105	CCATGAGTCTCGGCCCAAAACCCCTACTAACTGTTGGATATGCTTCCCGCTGAACTTCAR	6164	
Db	571	CCATGAGTCTCGGCCCAAAACCCCTACTAACTGTTGGATATGCTTCCCGCTGAACTTCAG	630	
QY	6165	GCCATATGTTTCAATCCCTGTACCTCGAAACAAATGGAAACAATTCAGCAGAGAAATAAACAC	6224	
Db	631	GCCATATGTTTCAATCCCTGTACCTCGAAACAAATGGAAACAATTCAGCAGAGAAATAAACAC	690	
QY	6225	CACTTCCGTTTTAGTAGGACCTCTTTGTTTCCAATSTGGAAATAAACCCATACCTCAAACT	6284	
Db	691	CACTTCCGTTTTAGTAGGACCTCTTTGTTTCCAATCTGGAAATAAACCCATACCTCAAACT	750	

QY 6285 CACCTGTGTAAATTTAGCAATACTACTACATACACCAACTTCCCAATGCATCAGTGGGT 6344  
 |||||  
 Db 751 CACCTGTGTAAATTTAGCAATACTACTACATACACCAACTTCCCAATGCATCAGTGGGT 810  
 |||||  
 QY 6345 AACTCTCTCCACACAAATAGTCTGCCTACCCTCAGGA 6381  
 |||||  
 Db 811 AACTCTCTCCACACAAATAGTCTGCCTACCCTCAGGA 847  
 |||||

RESULT 9  
 BX408733 883 bp mRNA linear EST 30-APR-2004  
 LOCUS BX408733 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE012U24  
 DEFINITION 5-PRIME, mRNA sequence.

ACCESSION BX408733  
 VERSION BX408733.1 GI:30635957  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 883)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.  
 This sequence belongs to sequence cluster 4215.r  
 For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?s=CS0BAA008ZE02\_CS00686\_1&c=4215.r

FEATURES  
 source  
 1..883  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DE012U24"  
 /tissue type="PLACENTA"  
 /clone\_lib="Homo sapiens PLACENTA"  
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN  
 Query Match 5.8%; Score 437; DB 4; Length 883;  
 Best Local Similarity 99.4%; Pred. No. 2.2e-232;  
 Matches 637; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGAGAAATGAGCGTCCCGAAATATTTGATGCCCATCTGTATGAGGAGTCT 5744  
 |||||  
 Db 75 AGAGTTTCTATGAGAAATGAGCGTCCCGAAATATTTGATGCCCATCTGTATGAGGAGTCT 134  
 |||||

QY 5745 TTSTAAGGGAACCCCACTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 5804  
 |||||  
 Db 135 TTCTAAGGGAACCCCACTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 194  
 |||||

QY 5805 TGGCACTCTTTGATGCATGCAAAATCTATTTGGACAGGAAAATGATTAATCCTAG 5864  
 |||||  
 Db 195 TGGCACTCTTTGATGCATGCAAAATCTATTTGGACAGGAAAATGATTAATCCTAG 254  
 |||||

QY 5865 TTGTCTCGGAGGACTTGGAGTCACTCTGTGTGGACTTACTTCAACCAAACTGGTATGTC 5924  
 |||||  
 Db 255 TTGTCTCGGAGGACTTGGAGTCACTCTGTGTGGACTTACTTCAACCAAACTGGTATGTC 314  
 |||||

QY 5925 TGATGGGGGTGGAGTTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAGTAAATCTCCCA 5984  
 |||||  
 Db 315 TGATGGGGGTGGAGTTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAGTAAATCTCCCA 374  
 |||||  
 QY 5985 ACTCACCGGGGTACATGGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAAACTACA 6044  
 |||||  
 Db 375 ACTCACCGGGGTACATGGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAAACTACA 434  
 |||||  
 QY 6045 TGAACCCCTCGGTACCCATCTCGCTCGTGAAGCCCTATTTAATACCAACCCCTCACTGGGT 6104  
 |||||  
 Db 435 TGAACCCCTCGGTACCCATCTCGCTCGTGAAGCCCTATTTAATACCAACCCCTCACTGGGT 494  
 |||||  
 QY 6105 CATGAGTCTCGGCCCAAAACCTTACTAACTGTGGATATGCCCTCCCTCGAATCTCAR 6164  
 |||||  
 Db 495 CCATGAGGTCTCGGCCCAAAACCTTACTAACTGTGGATATGCCCTCCCTCGAATCTCAG 554  
 |||||  
 QY 6165 GCATATGTTTCAATCCCTGTACCTGGAACAACCTTCAGCACAGAAATAAACAC 6224  
 |||||  
 Db 555 GCATATGTTTCAATCCCTGTACCTGGAACAACCTTCAGCACAGAAATAAACAC 614  
 |||||  
 QY 6225 CACTTCCGTTTTAGTAGGACCTCTTGTTCCTCAATSTGGAATTAACCCATACCTCAAACT 6284  
 |||||  
 Db 615 CACTTCCGTTTTAGTAGGACCTCTTGTTCCTCAATSTGGAATTAACCCATACCTCAAACT 674  
 |||||  
 QY 6285 CACTGTGTAAATTTAGCAATACTACTACATACACCAACT 6325  
 |||||  
 Db 675 CACTGTGTAAATTTAGCAATACTACTACATACACCAACT 715  
 |||||

RESULT 10  
 BX326647/c 872 bp mRNA linear EST 07-APR-2004  
 LOCUS BX326647 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 DEFINITION clone CS0DI051YM13 3-PRIME, mRNA sequence.

ACCESSION BX326647  
 VERSION BX326647  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 872)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4215.r  
 For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?s=CS0BAD010ZC11\_AD00948\_1&c=4215.r

FEATURES  
 source  
 Location/Qualifiers  
 1..872  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DI051YM13"  
 /tissue type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN



Query Match 5.7%; Score 431; DB 4; Length 872;  
 Best Local Similarity 95.4%; Pred. No. 5.2e-229;  
 Matches 631; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6038 AACTACATGAACCCCTCCGTACCCATACCTCGCTGGTAAGCCTATTAAATACCAACCTCA 6097  
 DB 693 AACTACATGAACCCCTCCGTACCCATACCTCGCTGGTAAGCCTATTAAATACCAACCTCA 634

QY 6098 CTGGGCTCCATAGAGTCTCGGCCCAAAACCCCTACTAACTGTGGATGATGCTCCCTCTGA 6157  
 DB 633 CTGGGCTCCATAGAGTCTCGGCCCAAAACCCCTACTAACTGTGGATGATGCTCCCTCTGA 574

QY 6158 ACTTCARGCCATATGTTTCAATCCCTGCTGAACATGGAACCACTTCAGACAGAAA 6217  
 DB 573 ACTTCARGCCATATGTTTCAATCCCTGCTGAACATGGAACCACTTCAGACAGAAA 514

QY 6218 TAAACACCACTTCGCTTTTAAAGTGGACCTCTGTTTCCAAATSTGGAATTAACCCATACCT 6277  
 DB 513 TAAACACCACTTCGCTTTTAAAGTGGACCTCTGTTTCCAAATSTGGAATTAACCCATACCT 454

QY 6278 CAAACCTCAGCTGTGTAATAATTTAGCAATACATACACAACTCCCAATGCATCA 6337  
 DB 453 CAAACCTCAGCTGTGTAATAATTTAGCAATACATACACAACTCCCAATGCATCA 394

QY 6338 GGTGGGTAACTCTCCCAACAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTG 6397  
 DB 393 GGTGGGTAACTCTCCCAACAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTG 334

QY 6398 GTACCTCAGCTATCGTTGTTTGAATGGCTCTTCAGAACTATGTCCTCTCATCTCT 6457  
 DB 333 GTACCTCAGCTATCGTTGTTTGAATGGCTCTTCAGAACTATGTCCTCTCATCTCT 274

QY 6458 TAGTGGCCCCATATGRCATCTACACTGAACAAGATTTATACAGTTATGTCATATCAAGC 6517  
 DB 273 TAGTGGCCCCATATGRCATCTACACTGAACAAGATTTATACAGTTATGTCATATCAAGC 214

QY 6518 CCGCAACAAGAGTACCAATCTTCCTTTGTTATAGAGCAGGAGTCTAGGTGCAC 6577  
 DB 213 CCGCAACAAGAGTACCAATCTTCCTTTGTTATAGAGCAGGAGTCTAGGTGCAC 154

QY 6578 TAGGTACTGCAATGGCGGTATCACAACCTCTACTCAGTTCTACTCAAACTATCTCAAG 6637  
 DB 153 TAGGTACTGCAATGGCGGTATCACAACCTCTACTCAGTTCTACTCAAACTATCTCAAG 94

QY 6638 AACTAAATGGGACATGGAACGGGTGCGGACTCC 6672  
 DB 93 AACTAAATGGGACATGGAACGGGTGCGGACTCC 59

RESULT 11  
 BX389656  
 LOCUS  
 DEFINITION  
 BX389656 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 Clone CS0D1051YM13 5-PRIME, mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

BX389656 931 bp mRNA linear EST 29-APR-2004  
 BX389656 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 Clone CS0D1051YM13 5-PRIME, mRNA sequence.

EST.  
 BX389656.1 GI:30462930  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
 1 (bases 1 to 931)  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 4215.r

For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?cs=CS0BAF027ZB05\_AF02531\_1&c=4215.r

## FEATURES

source  
 Location/Qualifiers  
 1..931  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1051YM13"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /notes="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 5.4%; Score 411; DB 4; Length 931;  
 Best Local Similarity 99.3%; Pred. No. 9e-218;  
 Matches 611; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5853 GATTAAATCCTAGTGTGCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCA 5912  
 DB 7 GATTAAATCCTAGTGTGCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCA 66

QY 5913 AACTGTATGTCTGATGGGGTGGAGTTCAGATCAGGCAAGAGAAAACATGTAAAGA 5972  
 DB 67 AACTGTATGTCTGATGGGGTGGAGTTCAGATCAGGCAAGAGAAAACATGTAAAGA 126

QY 5973 AGTAATCTCCCACTCAACCGGTACATGCACCTCTAGCCCTACAAAGGACTAGATCT 6032  
 DB 127 AGTAATCTCCCACTCAACCGGTACATGCACCTCTAGCCCTACAAAGGACTAGATCT 186

QY 6033 CTCAAATACATGAACCCCTCCGTACCCATCTCGCTGGTAAGCCTATTTAATACCAC 6092  
 DB 187 CTCAAATACATGAACCCCTCCGTACCCATCTCGCTGGTAAGCCTATTTAATACCAC 246

QY 6093 CCTCACTGGGCTCCATGAGTCTCGGCCCAAAACCCCTACTAATCTGTGGATATGCTTCC 6152  
 DB 247 CCTCACTGGGCTCCATGAGTCTCGGCCCAAAACCCCTACTAATCTGTGGATATGCTTCC 306

QY 6153 CTGGAATCTCARGCCATATGTTTCAATCCCTGCTGACCAATGGAACCAACTTCAGCAC 6212  
 DB 307 CTGGAATCTCARGCCATATGTTTCAATCCCTGCTGACCAATGGAACCAACTTCAGCAC 366

QY 6213 AGAAATAAACACACCTTCCTGTTTGTAGTAGGACCTCTGTTTCCAATSTGGAATAACCCA 6272  
 DB 367 AGAAATAAACACACCTTCCTGTTTGTAGTAGGACCTCTGTTTCCAATCTGGAATAACCCA 426

QY 6273 TACCTCAAAACCTCACTGCTGTGTAATAATTTAGCAATACTACATACAAACCACTCCCAATG 6332  
 DB 427 TACCTCAAAACCTCACTGCTGTGTAATAATTTAGCAATACTACATACAAACCACTCCCAATG 486

QY 6333 CATCAGGTGGGTAACTCTCCCAACAATAGTCTGCTTACCTCAGGAATATTTTGT 6392  
 DB 487 CATCAGGTGGGTAACTCTCCCAACAATAGTCTGCTTACCTCAGGAATATTTTGT 546

QY 6393 CTGTGTACTCTCAGCTATGCTGTTGTAATGGCTCTTCCAGAACTCTATGCTTCTCTC 6452  
 DB 547 CTGTGTACTCTCAGCTATGCTGTTGTAATGGCTCTTCCAGAACTCTATGCTTCTCTC 606

QY 6453 ATTCTTAGTGCCTCC 6467  
 DB 607 ATTCTTAGTGCCTCC 621

RESULT 12  
 DA830020  
 LOCUS

792 bp mRNA linear EST 13-NOV-2005

DEFINITION DA830020 PLACE1 Homo sapiens cDNA clone PLACE1004263 5', mRNA  
sequence.  
ACCESSION DA830020  
VERSION DA830020.1 GI:82330398  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 792)  
AUTHORS Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,  
Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaquri, H.,  
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,  
Yoneyama, T., Otsuka, K., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,  
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,  
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.  
TITLE Diversification of Transcriptional Modulation: Large-scale  
Identification and Characterization of Putative Alternative  
Promoters of Human Genes  
JOURNAL Genome Res. 16 (1), 55-65 (2006)  
PUBMED 16344560  
COMMENT Contact: Takao Isogai  
FLJ Project (HRI Team)  
Helix Research Institute  
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: flj-cdna@nifty.com  
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix  
Research Institute (HRI); cDNA library construction: Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
HRI.  
FEATURES  
source Location/Qualifiers  
1..792  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="PLACE1004263"  
/tissue\_type="placenta"  
/clone\_lib="PLACE1"  
/note="Vector: pME18SFL3"  
ORIGIN  
Query Match 5.3%; Score 403; DB 9; Length 792;  
Best Local Similarity 99.5%; Pred. No. 2.8e-213;  
Matches 553; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 5685 AGAGTTTCTATGGAGATCGAGGTCCTCCGGAATATTGATGCCCATCGTATAGGAGTCT 5744  
DB 84 AGAGTTTCTATGGAGATCGAGGTCCTCCGGAATATTGATGCCCATCGTATAGGAGTCT 143  
QY 5745 TTSTAAGGAACCCACCTTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTC 5804  
DB 144 TTCTAAGGAGACCCACCTTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTC 203  
QY 5805 TGCCACTCTTTCATGCATGCAAAATCTCATTTATGGACAGGAAAATGATTAATCCTTAG 5864  
DB 204 TGCCACTCTTTCATGCATGCAAAATCTCATTTATGGACAGGAAAATGATTAATCCTTAG 263  
QY 5865 TTGTCTCTGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAACTGGTATGTC 5924  
DB 264 TTGTCTCTGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCACCCAACTGGTATGTC 323  
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAGAAAGATTAATCTCCA 5984  
DB 324 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAGAAAGATTAATCTCCA 383  
QY 5985 ACTCACCGGGTACATGGCACTCTAGGCCCTCAAAAGGACTAGATCTCTCAAAACTACA 6044  
DB 384 ACTCACCGGGTACATGGCACTCTAGGCCCTCAAAAGGACTAGATCTCTCAAAACTACA 443  
QY 6045 TGAACCCCTCCGTACCCTACTCGCTGGTAAAGCCTATTATTAACACCCTCCTACCTGGGCT 6104

Db 444 TGAACCCCTCGTACCCTACTCGCTGGTAAAGCCTATTATTAACACCCTCCTACCTGGGCT 503  
QY 6105 CCATGAGTCTCGGCCCAAAACCCCTACTAATCTGTTGGATATGCTCCGCCCTGAATTCAR 6164  
Db 504 CCATGAGTCTCGGCCCAAAACCCCTACTAATCTGTTGGATATGCTCCGCCCTGAATTCAG 563  
QY 6165 GCCATATGTTTCAATCCCTGTACCTGACCAATGGAACTTCAGCAGAGAAATAAACAC 6224  
Db 564 GCCATATGTTTCAATCCCTGTACCTGACCAATGGAACTTCAGCAGAGAAATAAACAC 623  
QY 6225 CACTTCCGTTTATAGTA 6240  
Db 624 CACTTCCGTTTATAGTA 639  
RESULT 13  
LOCUS DA836641 PLACE6 Homo sapiens cDNA clone PLACE6002347 5', mRNA  
DEFINITION DA836641  
ACCESSION DA836641  
VERSION DA836641.1 GI:81213784  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 564)  
AUTHORS Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,  
Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaquri, H.,  
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,  
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,  
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,  
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.  
TITLE Diversification of Transcriptional Modulation: Large-scale  
Identification and Characterization of Putative Alternative  
Promoters of Human Genes  
JOURNAL Genome Res. 16 (1), 55-65 (2006)  
PUBMED 16344560  
COMMENT Contact: Takao Isogai  
FLJ Project (HRI Team)  
Helix Research Institute  
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: flj-cdna@nifty.com  
NEO human cDNA project (New Energy and Industrial Technology  
Developmental Organization, Japan); cDNA library construction:  
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,  
Research Association for Biotechnology (RAB) and Biotechnology  
Center, National Institute of Technology and Evaluation; 3'-end one  
pass sequencing: RAB.  
FEATURES  
source Location/Qualifiers  
1..564  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="PLACE6002347"  
/tissue\_type="placenta"  
/clone\_lib="PLACE6"  
/note="Vector: pME18SFL3"  
ORIGIN  
Query Match 5.0%; Score 379; DB 9; Length 564;  
Best Local Similarity 99.4%; Pred. No. 8.2e-200;  
Matches 529; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 5812 CTTTGCATGCATGCAAAATCTCATTTATGGACAGGAAAATGATTAATCCTTAGTTCCT 5871  
Db 1 CTTTGCATGCATGCAAAATCTCATTTATGGACAGGAAAATGATTAATCCTTAGTTCCT 60

5872 GGAGGACTTGGAGTCTGCTGTTGGACTTACTTCCACCAAACTGGTATGTTGATGG 5931  
 |||||  
 Db 61 GGAGGACTTGGAGTCTGCTGTTGGACTTACTTCCACCAAACTGGTATGTTGATGG 120  
 |||||  
 QY 5932 GTGGGATTCAGATCAGGCAAGAGAAAAACATGTAAAGAACTGTCTCCCAACTCACC 5991  
 |||||  
 Db 121 GTGGGATTCAGATCAGGCAAGAGAAAAACATGTAAAGAACTGTCTCCCAACTCACC 180  
 |||||  
 QY 5992 SGGGTACATGGCACTTAGCCCTTACCAAGGACTAGATCTCTCAAACTACATGAAC 6051  
 |||||  
 Db 181 CGGGTACATGGCACTTAGCCCTTACCAAGGACTAGATCTCTCAAACTACATGAAC 240  
 |||||  
 QY 6052 CTCGGTACCATACTCGCCCTGTAGCCTATTTAATACCAACCTCTCACTGGGCTCCATGAG 6111  
 |||||  
 Db 241 CTCGGTACCATACTCGCCCTGTAGCCTATTTAATACCAACCTCTCACTGGGCTCCATGAG 300  
 |||||  
 QY 6112 GTCTCGGCCCAAAACCTTACTAACTGTGGATATGCTTCCCTCGAATTTCAAGCCATAT 6171  
 |||||  
 Db 301 GTCTCGGCCCAAAACCTTACTAACTGTGGATATGCTTCCCTCGAATTTCAAGCCATAT 360  
 |||||  
 QY 6172 GTTTCATCCCTGTACTGAACATGGAACCACTTACGACAGAAATAAACCACTTCC 6231  
 |||||  
 Db 361 GTTTCATCCCTGTACTGAACATGGAACCACTTACGACAGAAATAAACCACTTCC 420  
 |||||  
 QY 6232 GTTTTAGTAGGACCTCTGTTTCCAACTSTGGAATAACCACTTCAACCTCACTCT 6291  
 |||||  
 Db 421 GTTTTAGTAGGACCTCTGTTTCCAACTSTGGAATAACCACTTCAACCTCACTCT 480  
 |||||  
 QY 6292 GTAAATTTAGCAATACTACATACACAACTCCCAATGATCAGGTGGG 6343  
 |||||  
 Db 481 GTAAATTTAGCAATACTACATACACAACTCCCAATGATCAGGTGGG 532  
 |||||

RESULT 14  
 BX347314  
 LOCUS  
 DEFINITION BX347314 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 clone CS0D1044YK06 5-PRIME, mRNA sequence.  
 ACCESSION BX347314  
 VERSION BX347314.1 GI:30365192  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4215.r

For more information about this cluster, see  
[http://www.genoscope.cns.fr/cdna?&CS0BAA016ZH06\\_CS01489\\_1&c=4215.r](http://www.genoscope.cns.fr/cdna?&CS0BAA016ZH06_CS01489_1&c=4215.r)

FEATURES  
 source  
 1..924  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1044YK06"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was

digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
 Query Match 5.0%; Score 377; DB 4; Length 924;  
 Best Local Similarity 99.6%; Pred. No. 1.1e-198;  
 Matches 477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 5685 AGAGTTTCTATGGAGAAATCGAGCGTCCCGGAAATATTGATGCCCACTCGTATAGGAGTCT 5744  
 |||||  
 Db 216 AGAGTTTCTATGGAGAAATCGAGCGTCCCGGAAATATTGATGCCCACTCGTATAGGAGTCT 275  
 |||||  
 QY 5745 TTSTAAGGAAACCCCACTTCACTGCCCAACCCATATGCCCCGCAACTGCTATCACTC 5804  
 |||||  
 Db 276 TTSTAAGGAAACCCCACTTCACTGCCCAACCCATATGCCCCGCAACTGCTATCACTC 335  
 |||||  
 QY 5805 TGCCACTCTTTTCATGTCATGCAAACTACTATTATTGGACAGAAATAATGATTAATCCTAG 5864  
 |||||  
 Db 336 TGCCACTCTTTTCATGTCATGCAAACTACTATTATTGGACAGAAATAATGATTAATCCTAG 395  
 |||||  
 QY 5865 TTGTCTCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCAACTGGTATGTC 5924  
 |||||  
 Db 396 TTGTCTCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCAACTGGTATGTC 455  
 |||||  
 QY 5925 TGATGGGGTGGAGTTCAGATCAGGCAAGAGAAAAAATGTAAGAAAGTAAATCTCCCA 5984  
 |||||  
 Db 456 TGATGGGGTGGAGTTCAGATCAGGCAAGAGAAAAAATGTAAGAAAGTAAATCTCCCA 515  
 |||||  
 QY 5985 ACTCACCCGGGTACATGGCACCCTTAGCCCTTACAAAGGACTAGATCTCTCAAACTACA 6044  
 |||||  
 Db 516 ACTCACCCGGGTACATGGCACCCTTAGCCCTTACAAAGGACTAGATCTCTCAAACTACA 575  
 |||||  
 QY 6045 TGAAACCCCTCCGTACCCATCTCGCTCGTGAAGCCCTATTATATACACCCTCACTGGGT 6104  
 |||||  
 Db 576 TGAAACCCCTCCGTACCCATCTCGCTCGTGAAGCCCTATTATATACACCCTCACTGGGT 635  
 |||||  
 QY 6105 CCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTTGGATATGCTTCCCTCGAATTTCA 6163  
 |||||  
 Db 636 CCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTTGGATATGCTTCCCTCGAATTTCA 694  
 |||||

RESULT 15  
 DA838045  
 LOCUS  
 DEFINITION DA838045 PLACE6 Homo sapiens cDNA clone PLACE6004226 5', mRNA  
 sequence.  
 ACCESSION DA838045  
 VERSION DA838045.1 GI:81276957  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.  
 1 (bases 1 to 538)  
 Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,  
 Yanahita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakeguri, H.,  
 Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kishida, N.,  
 Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,  
 Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,  
 Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.  
 Identification of Transcriptional Modulation: Large-scale  
 Promoters of Human Genes

JOURNAL  
 PUBMED  
 COMMENT  
 Contact: Takao Isogai  
 FLJ Project (HRI Team)  
 Helix Research Institute  
 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan  
 Tel: 81-438-52-3975  
 Fax: 81-438-52-3986  
 Email: flj-cdna@nifty.com

Genome Res. 16 (1), 55-65 (2006)  
 16344560

NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.

FEATURES  
source  
1..538  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="PLACE6004226"  
/tissue\_type="placenta"  
/clone\_lib="PLACE6"  
/note="Vector: pME18SFL3"

ORIGIN

Query Match 4.9%; Score 374; DB 9; Length 538;  
Best Local Similarity 99.6%; Pred. No. 5.3e-197;  
Matches 474; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 5688 GTTCTATGGAGAAATGCGAGCTCCCGGAAATATTCATGCCCATCGTATAGGAGTCTTTS 5747  
Db 1 GTTCTATGGAGAAATGCGAGCTCCCGGAAATATTCATGCCCATCGTATAGGAGTCTTTC 60  
QY 5748 TAAAGGAAACCCACCTTCACCTGCCACACCCATATGCCCGCAACTGCTATCACTCTGC 5807  
Db 61 TAAAGGAAACCCACCTTCACCTGCCACACCCATATGCCCGCAACTGCTATCACTCTGC 120  
QY 5808 CACTCTTTTCATGCATGCAATACCTATTATTTGGACAGAGAAATGATTAATCTAGTTG 5867  
Db 121 CACTCTTTTCATGCATGCAATACCTATTATTTGGACAGAGAAATGATTAATCTAGTTG 180  
QY 5868 TCCTGGAGGACTTGGAGTCACCTGTCTGTGTGGACTTACTTCAACCAAACTGGTATGCTGA 5927  
Db 181 TCCTGGAGGACTTGGAGTCACCTGTCTGTGTGGACTTACTTCAACCAAACTGGTATGCTGA 240  
QY 5928 TGGGGGTGGAGTCAAGATCAGGCAAGAGAGAAACATGTAAAGAGTAATCTCCCACT 5987  
Db 241 TGGGGGTGGAGTCAAGATCAGGCAAGAGAGAAACATGTAAAGAGTAATCTCCCACT 300  
QY 5988 CACCSGGGTACATGGCACCCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGA 6047  
Db 301 CACCSGGGTACATGGCACCCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGA 360  
QY 6048 AACCCCTCCGTACCCATCTCGCTGGTAAGCCCTATTTAATACCACCCCTCACTGGGCTCCA 6107  
Db 361 AACCCCTCCGTACCCATCTCGCTGGTAAGCCCTATTTAATACCACCCCTCACTGGGCTCCA 420  
QY 6108 TGAGGTCTGGCCCAAAACCCCTACTAACTGTGGATATGCCTCCCTGAACTTCA 6163  
Db 421 TGAGGTCTGGCCCAAAACCCCTACTAACTGTGGATATGCCTCCCTGAACTTCA 476

RESULT 16  
AUI138405  
LOCUS AUI138405 PLACE1 Homo sapiens cDNA clone PLACE1008489 5', mRNA  
DEFINITION sequence.  
ACCESSION AUI138405  
VERSION AUI138405.2 GI:55780428  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 721)  
AUTHORS Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T., Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H., Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N., Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M., Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,

Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.  
Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes  
Genome Res. 16 (1), 55-65 (2006)  
16344560  
On Oct 25, 2000 this sequence version replaced gi:10999926.  
Contact: Takao Isogai  
FLJ Project (HRI Team)  
Helix Research Institute  
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: flj-cdna@nifty.com  
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES  
source

1..721  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="PLACE1008489"  
/tissue\_type="placenta"  
/clone\_lib="PLACE1"  
/note="Vector: pME18SFL3"

ORIGIN

Query Match 4.9%; Score 373; DB 1; Length 721;  
Best Local Similarity 99.4%; Pred. No. 2e-196;  
Matches 523; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 6165 GCCATATGTTTCAATCCCTGTACCTGAACAAATGGAACAACTTCAGCAGAGAAATAAACAC 6224  
Db 23 GCCATATGTTTCAATCCCTGTACCTGAACAAATGGAACAACTTCAGCAGAGAAATAAACAC 82  
QY 6225 CACTTCGGTTTTAGTAGGACCTCTGTTCCTCAATSTGGAAATACCCATACCTCAAACT 6284  
Db 83 CACTTCGGTTTTAGTAGGACCTCTGTTCCTCAATSTGGAAATACCCATACCTCAAACT 142  
QY 6285 CACTGTGTAAATTTTAGCAATACCTACATACAAACAACTCCCAATGCATCAGGTGGGT 6344  
Db 143 CACTGTGTAAATTTTAGCAATACCTACATACAAACAACTCCCAATGCATCAGGTGGGT 202  
QY 6345 AACTCCTCCCAACAATAATGTCCTACCTACCTCAGGAATATTTTGTCTGTGTACCTC 6404  
Db 203 AACTCCTCCCAACAATAATGTCCTACCTACCTCAGGAATATTTTGTCTGTGTACCTC 262  
QY 6405 AGCCTATCGTTTGTGATGGCTCTTCAGATCTATGTGCTTCCTCTCATTTCTTAGTGCC 6464  
Db 263 AGCCTATCGTTTGTGATGGCTCTTCAGATCTATGTGCTTCCTCTCATTTCTTAGTGCC 322  
QY 6465 CCCATGRCCTATCTACCTGAAACAAGATTTATACAGTTATGTCTATATCTAAGCCCCGCAA 6524  
Db 323 CCCTATGACCATCTACCTGAAACAAGATTTATACAGTTATGTCTATATCTAAGCCCCGCAA 382  
QY 6525 CAAAAGAGTACCCATTTCTTCTTTTATAGGACAGAGTGTAGTGACATAGGTAC 6584  
Db 383 CAAAAGAGTACCCATTTCTTCTTTTATAGGACAGAGTGTAGTGACATAGGTAC 442  
QY 6585 TGGCATTTGGCGGTATCACAACTCTACTCAGTTCTTACTACAACTATCTCAGAACTAAA 6644  
Db 443 TGGCATTTGGCGGTATCACAACTCTACTCAGTTCTTACTACAACTATCTCAGAACTAAA 502  
QY 6645 TGGGACATGAAACGGGTCCGCACTCCCTGGTCACTTGAAGAT 6690  
Db 503 TGGGACATGAAACGGGTCCGCACTCCCTGGTCACTTGAAGAT 548

RESULT 17  
DA832031  
LOCUS DA832031 761 bp mRNA linear EST 13-NOV-2005

```

DEFINITION DA832031 PLACE1 Homo sapiens cDNA clone PLACE1008285 5', mRNA
sequence.
ACCESSION DA832031
VERSION DA832031.1 GI:82376942
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 761)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute (HRI); cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
HRI.

FEATURES             source
    source
        1..761
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="PLACE1008285"
            /tissue_type="placenta"
            /clone_lib="PLACE1"
            /note="Vector: pME18SFL3"

ORIGIN
Query Match          4.9%; Score 373; DB 9; Length 761;
Best Local Similarity 99.2%; Pred. No. 2e-196;
Matches 623; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5835 TTATTGGACAGGAAAATGATTAACTCTAGTTGCTCGGAGGACTTGGAGTCACTGTCTG 5894
DB 1 TTATTGGACAGGAAAATGATTAACTCTAGTTGCTCGGAGGACTTGGAGTCACTGTCTG 60

QY 5895 TTGACTTACTTCAACCAACTGGTATGCTGATGGGGTGGAGTTCAAGTACAGCAAG 5954
DB 61 TTGACTTACTTCAACCAACTGGTATGCTGATGGGGTGGAGTTCAAGTACAGCAAG 120

QY 5955 AGAAAAATGTTAAAGAACTAATCTCCCACTCACCAGGTTACATGGCACTCTAGCCCC 6014
DB 121 AGAAAAATGTTAAAGAACTAATCTCCCACTCACCAGGTTACATGGCACTCTAGCCCC 180

QY 6015 CTCAAAAGGACTAGATCTCTCAAAAATACATGAACACCTCCGTAACCATCTCGCTGGT 6074
DB 181 CTCAAAAGGACTAGATCTCTCAAAAATACATGAACACCTCCGTAACCATCTCGCTGGT 240

QY 6075 AAGCCTATTAAATACACCTCACTGGCTCCATGAGTCTCGGCCCAAAACCTACTAA 6134
DB 241 AAGCCTATTAAATACACCTCACTGGCTCCATGAGTCTCGGCCCAAAACCTACTAA 300

QY 6135 CTGTTGGATGCTCCCTCGAATCTCARGCCATATGTTTCAATCCTGTACTGAACA 6194
DB 301 CTGTTGGATGCTCCCTCGAATCTCARGCCATATGTTTCAATCCTGTACTGAACA 360

QY 6195 ATGGAACAACCTCAGCACAGAAAATAACCACTTCGTTTATAGGACCTCTTGTTC 6254

```

```

Db 361 ATGGAACAACCTCAGCACAGAAAATAGACACCACTCCGTTTATAGTACGCTCTTGTTC 420
QY 6255 CAATSTGGAATAAACCCATACCTCAAACTCAGCTGTGTAAATTTTAGCAATACTACATA 6314
Db 421 CAATCTGGAATAAACCCATACCTCAAACTCAGCTGTGTAAATTTTAGCAATACTACATA 480
QY 6315 CACAACCAACTCCCAATGCATCAGGTGGGTAATCTCTCCACACAAATAGTGTGCTTACC 6374
Db 481 CACAACCAACTCCCAATGCATCAGGTGGGTAATCTCTCTCCACACAAATAGTGTGCTTACC 540
QY 6375 CTCAGAATAATTTTGTCTGTGTGCTCAGCTCAGCTATCGTTGTGAATGGCTCTTACA 6434
Db 541 CTCAGGAATAATTTTGTCTGTGTGCTCAGCTCAGCTATCGTTGTGAATGGCTCTTACA 600
QY 6435 ATCTATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6462
Db 601 ATCTATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 628

RESULT 18
DA849357
LOCUS DA849357 PLACE6 Homo sapiens cDNA clone PLACE6018680 5', mRNA
DEFINITION sequence.
ACCESSION DA849357
VERSION DA849357.1 GI:82140231
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 540)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N.,
Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
Diversification of Transcriptional Modulation: Large-scale
Identification and Characterization of Putative Alternative
Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)
16344560
Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan); cDNA library construction:
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.

FEATURES             Location/Qualifiers
    source
        1..540
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="PLACE6018680"
            /tissue_type="placenta"
            /clone_lib="PLACE6"
            /note="Vector: pME18SFL3"

ORIGIN
Query Match          4.9%; Score 372; DB 9; Length 540;
Best Local Similarity 99.6%; Pred. No. 7e-196;
Matches 472; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```





1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.

This sequence belongs to sequence cluster 4215.x

For more information about this cluster, see  
[http://www.genoscope.cns.fr/cdna?S=CS0BAA008ZE02\\_CS00686\\_2&c=4215.x](http://www.genoscope.cns.fr/cdna?S=CS0BAA008ZE02_CS00686_2&c=4215.x)

#### FEATURES

Location/Qualifiers

1..921

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DF012XJ24"

/tissue\_type="PLACENTA"

/clone\_lib="Homo sapiens PLACENTA"

/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

#### ORIGIN

Query Match 4.9%; Score 369; DB 4; Length 921;

Best Local Similarity 99.6%; Pred. No. 3.5e-194;

Matches 469; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5748 TAAGGGAACCCCACTTCACTGCCACACCCCATATGCCCCGCAACTGCTATCACTCTGC 5807

Db 139 TAAGGGAACCCCACTTCACTGCCACACCCCATATGCCCCGCAACTGCTATCACTCTGC 198

QY 5808 CACTTTTGATGATGCAAAATCTCATTTATGGACAGGAAAAATGATTAATCCTAGTTG 5867

Db 199 CACTTTTGATGATGCAAAATCTCATTTATGGACAGGAAAAATGATTAATCCTAGTTG 258

QY 5868 TCCTGGAGGACTTGGAGTCACTCTGTTGGACTTACTTCAACCAACTGGTATGCTGA 5927

Db 259 TCCTGGAGGACTTGGAGTCACTCTGTTGGACTTACTTCAACCAACTGGTATGCTGA 318

QY 5928 TGGGGGTGGAGTTCAAGATCAGCAAGAGAAAAACATGTAAGAAAGTAACTCCCAACT 5987

Db 319 TGGGGGTGGAGTTCAAGATCAGCAAGAGAAAAACATGTAAGAAAGTAACTCCCAACT 378

QY 5988 CACCSGGGTACATGGCACTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGA 6047

Db 379 CACCSGGGTACATGGCACTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGA 438

QY 6048 AACCTCCGTACCATCTAGCCCTGTAAGCCTATTTAATACACCCCTCACTGGGCTCCA 6107

Db 439 AACCTCCGTACCATCTAGCCCTGTAAGCCTATTTAATACACCCCTCACTGGGCTCCA 498

QY 6108 TGAGGTCTCGGCCCAAAACCTACTAACTGTTGGATATGCTCCCTCAACTTTCARGCC 6167

Db 499 TGAGGTCTCGGCCCAAAACCTACTAACTGTTGGATATGCTCCCTCAACTTTCARGCC 558

QY 6168 ATATGTTTCAATCCCTGTACCTGAAACAATGGAAACAACCTTCAGCACAGAAAT 6218

Db 559 ATATGTTTCAATCCCTGTACCTGAAACAATGGAAACAACCTTCAGCACAGAAAT 609

#### RESULT 21

DA846799

LOCUS

DA846799 PLACE6 Homo sapiens cDNA clone PLACE6015377 5', mRNA

DEFINITION

sequence.

ACCESSION

DA846799.1

VERSION

GI:82041683

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

#### REFERENCE

AUTHORS

Hominidae; Homo.

1 (bases 1 to 571)

Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,

Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,

Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,

Yoneyama,T., Otsuka,B., Kanda,K., Yokoi,T., Kondo,H., Wagatsuna,M.,

Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,

Tanabe,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.

Diversification of Transcriptional Modulation: Large-scale

Identification and Characterization of Putative Alternative

Promoters of Human Genes

Genome Res. 16 (1), 55-65 (2006)

16344560

Contact: Takao Isogai

FLJ Project (HRI Team)

Helix Research Institute

2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan

Tel: 81-438-52-3975

Fax: 81-438-52-3986

Email: flj-cdna@nifty.com

NEDO human cDNA project (New Energy and Industrial Technology

Developmental Organization, Japan); cDNA library construction:

Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,

Research Association for Biotechnology (RAB) and Biotechnology

Center, National Institute of Technology and Evaluation; 3'-end one

pass sequencing: RAB.

Location/Qualifiers

1..571

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="PLACE6015377"

/tissue\_type="placenta"

/clone\_lib="PLACE6"

/note="Vector: pME18SFL3"

ORIGIN

Query Match 4.8%; Score 367; DB 9; Length 571;

Best Local Similarity 99.3%; Pred. No. 4.5e-193;

Matches 567; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6003 CACCTCTAGCCCCCTCAAGGACTAGATCTCTCAAAACTACATGAACCCCTCGTACCCA 6062

Db 1 CACCTCTAGCCCCCTCAAGGACTAGATCTCTCAAAACTACATGAACCCCTCGTACCCA 60

QY 6063 TACTGCGCTGTGAAGCTATTAAATACACACCTCTCACTGGGCTCCATGAGGTCTCGGCCCA 6122

Db 61 TACTGCGCTGTGAAGCTATTAAATACACACCTCTCACTGGGCTCCATGAGGTCTCGGCCCA 120

QY 6123 AAACCCCTACTAACTGTGGATATGCTCCCTCGAATTTTCAAGCCATATGTTCAATCCC 6182

Db 121 AAACCCCTACTAACTGTGGATATGCTCCCTCGAATTTTCAAGCCATATGTTCAATCCC 180

QY 6183 TGCTACTGAACAATGGAACAACCTTCAGCAGAGAAATAACACACTTCCTCGTTTGTAGG 6242

Db 181 TGCTACTGAACAATGGAACAACCTTCAGCAGAGAAATAACACACTTCCTCGTTTGTAGG 240

QY 6243 ACCTCTTGTTCCTCAATSTGGAATAACCCATACCTCAACCTCCTCGTGTAAATTTAG 6302

Db 241 ACCTCTTGTTCCTCAATSTGGAATAACCCATACCTCAACCTCCTCGTGTAAATTTAG 300

QY 6303 CAATACTACATACACAACTCCCAATGCACTGAGGTGGTAACTCTCCACACCAAAAT 6362

Db 301 CAATACTACATACACAACTCCCAATGCACTGAGGTGGTAACTCTCCACACCAAAAT 360

QY 6363 AGCTGCGCTACCTCAGGAATATTTTGTCTGTGGTACTCAGGCTATCGTTTGTAGG 6422

Db 361 AGCTGCGCTACCTCAGGAATATTTTGTCTGTGGTACTCAGGCTATCGTTTGTAGG 420

QY 6423 TGGCTCTTCAGAACTATGTGCTTCTCTCATTTCTAGTGGCCCCCYATGRCCTATCTACAC 6482

Db 421 TGGCTCTTCAGAACTATGTGCTTCTCTCATTTCTAGTGGCCCCCYATGRCCTATCTACAC 480

QY 6483 TGACCAAGATTTATACAGTTATGTCTATCTAAGCCCGCAACAAAAGAGTACCCATTCT 6542  
 |||||  
 Db 481 TGACCAAGATTTATACAGTTATGTCTATCTAAGCCCGCAACAAAAGAGTACCCATTCT 540  
 |||||  
 QY 6543 TCCTTTTGTATAGGACGAGGAGTGCTAGGT 6573  
 |||||  
 Db 541 TCCTTTTGTATAGGACGAGGAGTGCTAGGT 571  
 |||||

RESULT 22  
 DA835080  
 LOCUS DA835080 PLAC66 Homo sapiens cDNA clone PLACE6000237 5', mRNA  
 DEFINITION sequence.  
 ACCESSION DA835080  
 VERSION DA835080.1 GI:83059534  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominiidae; Homo.

REFERENCE 1 (bases 1 to 541)  
 AUTHORS Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,  
 Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,  
 Ishii, S., Sugiyama, T., Saio, K., Isono, Y., Irie, R., Kushida, N.,  
 Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,  
 Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,  
 Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.  
 TITLE Diversification of Transcriptional Modulation: Large-scale  
 Identification and Characterization of Putative Alternative  
 Promoters of Human Genes  
 JOURNAL Genome Res. 16 (1), 55-65 (2006)  
 PUBMED 16344560  
 COMMENT Contact: Takao Isogai  
 FLJ Project (HRI Team)  
 Helix Research Institute  
 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan  
 Tel: 81-438-52-3975  
 Fax: 81-438-52-3986  
 Email: flj-cdna@nifty.com

NEDO human cDNA project (New Energy and Industrial Technology  
 Developmental Organization, Japan); cDNA library construction:  
 Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,  
 Research Association for Biotechnology (RAB) and Biotechnology  
 Center, National Institute of Technology and Evaluation; 3'-end one  
 pass sequencing: RAB.

FEATURES  
 source  
 1..541  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="PLACE6000237"  
 /tissue\_type="placenta"  
 /clone\_lib="PLACE6"  
 /note="Vector: pME18SFL3"

ORIGIN  
 Query Match 4.8%; Score 363; DB 9; Length 541;  
 Best Local Similarity 99.6%; Pred. No. 8e-191;  
 Matches 463; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGAGAAATCAGGTCCTCGGAAATATTGATGCCCATCGTATAGGAGTCT 5744  
 |||||  
 Db 77 AGAGTTTCTATGAGAAATCAGGTCCTCGGAAATATTGATGCCCATCGTATAGGAGTCT 136  
 |||||  
 QY 5745 TTSTAAGGAACCCCACTTCACTGCCCAACCCCATATGCCCGCAACTGCTATCACTC 5804  
 |||||  
 Db 137 TTSTAAGGAACCCCACTTCACTGCCCAACCCCATATGCCCGCAACTGCTATCACTC 196  
 |||||  
 QY 5805 TGCCACTCTTTGCATGCGCAAAATCACTATTATGGACAGGAAATGATTAATCCCTAG 5864  
 |||||  
 Db 197 TGCCACTCTTTGCATGCGCAAAATCACTATTATGGACAGGAAATGATTAATCCCTAG 256  
 |||||

QY 5865 TTGTCTCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCACCCAAACTGGTATGTC 5924  
 |||||  
 Db 257 TTGTCTCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCACCCAAACTGGTATGTC 316  
 |||||  
 QY 5925 TGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAAAGTAATCTCCCA 5984  
 |||||  
 Db 317 TGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAAAGTAATCTCCCA 376  
 |||||  
 QY 5985 ACTCACCGGGTACATGGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACA 6044  
 |||||  
 Db 377 ACTCACCGGGTACATGGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACA 436  
 |||||  
 QY 6045 TGAACCCCTCGTACCCCATACTCGCTGGTAAAGCTATTAAATACCCCTCAGTGGCT 6104  
 |||||  
 Db 437 TGAACCCCTCGTACCCCATACTCGCTGGTAAAGCTATTAAATACCCCTCAGTGGCT 496  
 |||||  
 QY 6105 CCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTTGGATATGCGCT 6149  
 |||||  
 Db 497 CCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTTGGATATGCGCT 541  
 |||||

RESULT 23  
 BX368078/c  
 LOCUS BX368078 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 DEFINITION clone CSODI051YM13 3-PRIME, mRNA sequence.

ACCESSION BX368078  
 VERSION BX368078.2 GI:46572908  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominiidae; Homo.

REFERENCE 1 (bases 1 to 846)  
 AUTHORS Li, W. B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT On May 8, 2003 this sequence version replaced gi:30447710.

Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 4215.r

For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?s=CS0AU0122D05\_U01092\_1&c=4215.r.

FEATURES  
 source  
 1..846  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CSODI051YM13"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
 Query Match 4.7%; Score 360; DB 4; Length 846;  
 Best Local Similarity 99.6%; Pred. No. 4e-189;  
 Matches 460; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 6260 TGGAAATAACCCATACCTCAAAACCTCAGCTGTGTAAAAATTTAGCAATACATACACAA 6319  
 |||||  
 Db 629 TGGAAATAACCCATACCTCAAAACCTCAGCTGTGTAAAAATTTAGCAATACATACACAA 570  
 |||||

```
QY 6320 CCACTCCCAATGATCAGTGGTAACTCTCTCCACACAAATAGTCTGCTACCTCAG 6379
|
|
|
Db 569 CCAACTCCCAATGATCAGTGGTAACTCTCTCCACACAAATAGTCTGCTACCTCAG 510
|
|
|
QY 6380 GAATATATTTTGTCTGTGGTACCTCAGCCTATCGTTGTTGAATGGCTCTTCAGAATCTA 6439
|
|
|
Db 509 GAATATATTTTGTCTGTGGTACCTCAGCCTATCGTTGTTGAATGGCTCTTCAGAATCTA 450
|
|
|
QY 6440 TGTGCTCTCTCATTTCTTAGTGGCCCCYATGRCATCTACATGAACAAGATTTATACA 6499
|
|
|
Db 449 TGTGCTCTCTCATTTCTTAGTGGCCCCATGACCATCTACATGAACAAGATTTATACA 390
|
|
|
QY 6500 GTTATGTCATATCTAAGCCCGCAACAAAGAGTACCATTCTCTCTTTTGTATAGGAG 6559
|
|
|
Db 389 GTTATGTCATATCTAAGCCCGCAACAAAGAGTACCATTCTCTCTTTTGTATAGGAG 330
|
|
|
QY 6560 CAGGAGTGTAGTGGTACCTAGTACTGGCATTTGGCGGTATCACAACTCTACTCAGTTCT 6619
|
|
|
Db 329 CAGGAGTGTAGTGGTACCTAGTACTGGCATTTGGCGGTATCACAACTCTACTCAGTTCT 270
|
|
|
QY 6620 ACTACAAATATCTCAAGAACTAAATGGGGACATGGAACGGGTGCGGACTCCCTGGTCA 6679
|
|
|
Db 269 ACTACAAATATCTCAAGAACTAAATGGGGACATGGAACGGGTGCGGACTCCCTGGTCA 210
|
|
|
QY 6680 CTTTGAAGATCACTTAACCTCCTAGCAGCAGTGTCTTC 6721
|
|
|
Db 209 CTTTGAAGATCACTTAACCTCCTAGCAGCAGTGTCTTC 168
|
|
|
```

## RESULT 24

```
DA845887
LOCUS
DEFINITION
  DA845887 PLAC66 Homo sapiens cDNA clone PLAC66014209 5', mRNA
sequence.
```

```
ACCESSION
  DA845887
```

```
VERSION
  DA845887.1 GI:81473746
```

```
KEYWORDS
  EST.
```

```
SOURCE
  Homo sapiens (human)
```

## ORGANISM

```

  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
```

```

  1 (bases 1 to 547)
```

```
REFERENCE
  Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
  Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
  Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kishida, N.,
  Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
  Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
  Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
  Diversification of Transcriptional Modulation: Large-scale
  Identification and Characterization of Putative Alternative
  Promoters of Human Genes
```

```
Genome Res. 16 (1), 55-65 (2006)
```

```
16344560
```

```
Contact: Takao Isogai
```

```
FLJ Project (HRI Team)
```

```
Helix Research Institute
```

```
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
```

```
Tel: 81-438-52-3975
```

```
Fax: 81-438-52-3986
```

```
Email: flj-cdna@nifty.com
```

```
NEDO human cDNA project (New Energy and Industrial Technology
Developmental Organization, Japan): cDNA library construction.
Helix Research Institute (HRI): 5'-end one pass sequencing: HRI,
Research Association for Biotechnology (RAB) and Biotechnology
Center, National Institute of Technology and Evaluation; 3'-end one
pass sequencing: RAB.
```

## FEATURES

```
source
```

```
1..547
```

```
/organism="Homo sapiens"
```

```
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"
```

```
/clone="PLAC66014209"
/tissue_type="placenta"
/clone_lib="PLAC66"
/notes="Vector: pME18SFL3"
```

## ORIGIN

```

Query Match      4.7%; Score 358; DB 9; Length 547;
Best Local Similarity 99.4%; Pred. No. 5.1e-188;
Matches 508; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 5748 TAAGGGAACCCCCACCTTCACTGCCACACCCCATATGCCCCGCAACTGCTATCACTCTGC 5807
```

```
Db 35 TAAGGGAACCCCCACCTTCACTGCCACACCCCATATGCCCCGCAACTGCTATCACTCTGC 94
```

```
QY 5808 CACTCTTTGATGATGCAATGCAATATCTATTATGGACAGGAAAAATGATTAATCTAGTTG 5867
```

```
Db 95 CACTCTTTGATGATGCAATGCAATATCTATTATGGACAGGAAAAATGATTAATCTAGTTG 154
```

```
QY 5868 TCTTGAGAGACCTTGGAGTCACTGTCTGTTGGACTTACTTCAACCAACTGGTATGTCGA 5927
```

```
Db 155 TCTTGAGAGACCTTGGAGTCACTGTCTGTTGGACTTACTTCAACCAACTGGTATGTCGA 214
```

```
QY 5928 TGGGGGTGGAGTTCAAGATCAGGCAAGAAAAACATGTAAPAGAGATTAATCTCCCAACT 5987
```

```
Db 215 TGGGGGTGGAGTTCAAGATCAGGCAAGAAAAACATGTAAPAGAGATTAATCTCCCAACT 274
```

```
QY 5988 CACCSGGGTACATGGCACCTCTAGCCCCCTACAAGGACTAGATCTCTCAAAACTACATGA 6047
```

```
Db 275 CACCSGGGTACATGGCACCTCTAGCCCCCTACAAGGACTAGATCTCTCAAAACTACATGA 334
```

```
QY 6048 AACCTCTCCGTACCCATCTGCTGTGAAGCCTATTATTAACACCTCACTGGGTCCA 6107
```

```
Db 335 AACCTCTCCGTACCCATCTGCTGTGAAGCCTATTATTAACACCTCACTGGGTCCA 394
```

```
QY 6108 TGAGGTCTCGGCCCAAAACCTCTACTAATCTGTTGGATATGCTCCCTGCACTTCAGGCC 6167
```

```
Db 395 CGAGGTCTCGGCCCAAAACCTCTACTAATCTGTTGGATATGCTCCCTGCACTTCAGGCC 454
```

```
QY 6168 ATATGTTTCAATCCCTGTACTGAAACAATGGAACAACCTTTCAGACAGAAATAACACCAC 6227
```

```
Db 455 ATATGTTTCAATCCCTGTACTGAAACAATGGAACAACCTTTCAGACAGAAATAACACCAC 514
```

```
QY 6228 TTCCGTTTTAGTAGGACCTCTTGTTCCTCAAT 6258
```

```
Db 515 TTCCGTTTTAGTAGGACCTCTTGTTCCTCAAT 545
```

```
RESULT 25
```

```
DA846543
```

```
LOCUS
```

```
DEFINITION
```

```
  DA846543 PLAC66 Homo sapiens cDNA clone PLAC66015035 5', mRNA
```

```
sequence.
```

```
ACCESSION
```

```
  DA846543
```

```
VERSION
```

```
  DA846543.1 GI:81817963
```

```
KEYWORDS
```

```
  EST.
```

```
SOURCE
```

```
  Homo sapiens (human)
```

```
ORGANISM
```

```

  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
```

```
  1 (bases 1 to 571)
```

```
REFERENCE
```

```
  Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T.,
```

```
  Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H.,
```

```
  Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kishida, N.,
```

```
  Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M.,
```

```
  Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A.,
```

```
  Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S.
```

```
  Diversification of Transcriptional Modulation: Large-scale
```

```
  Identification and Characterization of Putative Alternative
```

```
  Promoters of Human Genes
```

```
Genome Res. 16 (1), 55-65 (2006)
```

```
16344560
```

```
Contact: Takao Isogai
```

```
FLJ Project (HRI Team)
```

```
Helix Research Institute
```

```
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
```

```
Tel: 81-438-52-3975
```

```
Fax: 81-438-52-3986
```

```
Email: flj-cdna@nifty.com
```

```
NEDO human cDNA project (New Energy and Industrial Technology
```

```
Developmental Organization, Japan): cDNA library construction.
```

```
Helix Research Institute (HRI): 5'-end one pass sequencing: HRI,
```

```
Research Association for Biotechnology (RAB) and Biotechnology
```

```
Center, National Institute of Technology and Evaluation; 3'-end one
```

```
pass sequencing: RAB.
```

```
Location/Qualifiers
```

```
1..547
```

```
/organism="Homo sapiens"
```

```
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"
```

```
Contact: Takao Isogai
```

FLJ Project (HRI Team)  
Helix Research Institute  
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: flj-cdna@nifty.com

NEDO human cDNA project (New Energy and Industrial Technology  
Developmental Organization, Japan); cDNA library construction: HRI,  
Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,  
Research Association for Biotechnology (RAB) and Biotechnology  
Center, National Institute of Technology and Evaluation; 3'-end one  
pass sequencing: RAB.

FEATURES  
source  
Location/Qualifiers  
1..571  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="PLACE6015035"  
/tissue\_type="placenta"  
/clone\_lib="PLACE6"  
/note="Vector: pME18SPL3"

ORIGIN  
Query Match 4.7%; Score 357; DB 9; Length 571;  
Best Local Similarity 99.6%; Pred. No. 1.9e-187;  
Matches 457; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 5685 AGAGTTTCTATGAGAAATGACGCTCCGGAAATATGATGCCCATCGTATAGGAGTCT 5744  
Db 113 AGAGTTTCTATGAGAAATGACGCTCCGGAAATATGATGCCCATCGTATAGGAGTCT 172  
QY 5745 TTSTTAAGGAACCCACCTTCACTGCCCCACACCCATATGCCCGGAACCTGCTATCACTC 5804  
Db 173 TTCTAAGGAACCCACCTTCACTGCCCCACACCCATATGCCCGGAACCTGCTATCACTC 232  
QY 5805 TGCCACTCTTTCATGCAATGCAATCTCACTGCCCCACACCCATATGCCCGGAACCTGCTATCACTC 5864  
Db 233 TGCCACTCTTTCATGCAATGCAATCTCACTGCCCCACACCCATATGCCCGGAACCTGCTATCACTC 292  
QY 5865 TTGCTCTGGAGGACTTGGAGTCACTGCTCTGTTGGACTTACTTCAACCCAACTGGTATGTC 5924  
Db 293 TTGCTCTGGAGGACTTGGAGTCACTGCTCTGTTGGACTTACTTCAACCCAACTGGTATGTC 352  
QY 5925 TGATGGGGTGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAGAAGTAATCTCCCA 5984  
Db 353 TGATGGGGTGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAGAAGTAATCTCCCA 412  
QY 5985 ACTCACCGGGTACATGGCACCCTCTAGCCCTACAAAGGACTAGATCTCTCAAAACTACA 6044  
Db 413 ACTCACCGGGTACATGGCACCCTCTAGCCCTACAAAGGACTAGATCTCTCAAAACTACA 472  
QY 6045 TGAACCCCTCCGTACCATCTCGCTGTAAGCCCTATTATATACCAACCCCTCACTGGGCT 6104  
Db 473 TGAACCCCTCCGTACCATCTCGCTGTAAGCCCTATTATATACCAACCCCTCACTGGGCT 532  
QY 6105 CCATGAGGTCCTGGGCCAAACCCCTACTAACTGTTGGAT 6143  
Db 533 CCATGAGGTCCTGGGCCAAACCCCTACTAACTGTTGGAT 571

RESULT 26  
BX389657  
LOCUS  
DEFINITION  
BX389657 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
clone CS0D1051YM13 5-PRIME, mRNA sequence.

ACCESSION  
BX389657  
VERSION  
BX389657.2 GI:46975146  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1. (bases 1 to 995)  
Li.W.B., Gruber.C., Jessee.J. and Polayes.D.  
Full-length cDNA libraries and normalization  
unpublished (2001)  
On May 8, 2003 this sequence version replaced gi:30462931.

Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqre@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
4215.1

For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?s=CS0BAF027ZB05\_AFO2531\_2&c=4215.1

PEATUES  
source  
Location/Qualifiers

1..995  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1051YM13"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 4.7%; Score 356; DB 4; Length 995;  
Best Local Similarity 99.3%; Pred. No. 7.1e-187;  
Matches 556; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 5887 ACTGCTCTTGGAGTCTTACTTCCACCAACTGGTATGCTGATGGGGTGGAGTTCAGAT 5946  
Db 41 ACTGCTCTTGGAGTCTTACTTCCACCAACTGGTATGCTGATGGGGTGGAGTTCAGAT 100  
QY 5947 CAGGCAAGAGAAAAACATGTAAAGAGTAATCTCCCAACTCACCAGGATACATGGCACC 6006  
Db 101 CAGGCAAGAGAAAAACATGTAAAGAGTAATCTCCCAACTCACCAGGATACATGGCACC 160  
QY 6007 TCTAGCCCTCAAGAGCTAGATCTCTCAAACTACATGAAACCTCGTACCCATCT 6066  
Db 161 TCTAGCCCTCAAGAGCTAGATCTCTCAAACTACATGAAACCTCGTACCCATCT 220  
QY 6067 CGCCTGGTAAGCCTATTTAATACACCTCTAGGGCTCCATGAGGTCTCGGCCCAAAAC 6126  
Db 221 CGCCTGGTAAGCCTATTTAATACACCTCTAGGGCTCCATGAGGTCTCGGCCCAAAAC 280  
QY 6127 CTTACTAATCTTGGATATGCTCCCTCGTGAACCTTCARGCCATATGTTTCAATCCCTGTA 6186  
Db 281 CTTACTAATCTTGGATATGCTCCCTCGTGAACCTTCARGCCATATGTTTCAATCCCTGTA 340  
QY 6187 CTTGAACATGGAACAACCTTCAGCAGAGAAATAAACAACAATTCCTAGGACCT 6246  
Db 341 CTTGAACATGGAACAACCTTCAGCAGAGAAATAAACAACAATTCCTAGGACCT 400  
QY 6247 CTTGTTTCCCAATSTGAAATAACCATACCTCAACCTCTACCTGTGTAAATTTAGCAAT 6306  
Db 401 CTTGTTTCCCAATCTGGAATAACCATACCTCAACCTCTACCTGTGTAAATTTAGCAAT 460  
QY 6307 ACTCATACACAAACCAACTCCCAATGCATCAGGTGGTAACTCTCCCAACACAAATAGTC 6366  
Db 461 ACTCATACACAAACCAACTCCCAATGCATCAGGTGGTAACTCTCCCAACACAAATAGTC 520  
QY 6367 TGCCCTACCTCAGGAATATTTTCTCTGTGTGTACCTGAGCTATCGTTGTTGATGGC 6426  
Db 521 TGCCCTACCTCAGGAATATTTTCTCTGTGTGTACCTGAGCTATCGTTGTTGATGGC 580  
QY 6427 TCTTCAGAACTCTATGTGCTT 6446

```

Db      581  TCTTCAGAACTATGCTT 600
|||||
RESULT 27
BX430050/c      658 bp      mRNA      linear      EST 04-MAY-2004
LOCUS      BX430050 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE012Y024
DEFINITION      3-PRIME, mRNA sequence.
ACCESSION      BX430050
VERSION      BX430050.2 GI:47003301
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 658)
Li, W. B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 15, 2003 this sequence version replaced gi:30776872.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAB0022D11_CS00181_l&c=4215.r

FEATURES
source      Location/Qualifiers
1. .658
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DR012YJ24"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT_6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match      4.7%; Score 353; DB 4; Length 658;
Best Local Similarity 99.3%; Pred. No. 3.3e-185;
Matches 553; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      6165  GCCATATGTTTCAATCCCTGACTGAAACAATGGAAACAATTCCAGACAGAAATAAACAC 6224
|||||
Db      577  GCCATATGTTTCAATCCCTGACTGAAACAATGGAAACAATTCCAGACAGAAATAAACAC 518
|||||
QY      6225  CACTTCGGTTTATAGGACCTCTGTTCCATSTGGAAATACCCATACCTCAAACT 6284
|||||
Db      517  CACTTCGGTTTATAGGACCTCTGTTCCATSTGGAAATACCCATACCTCAAACT 458
|||||
QY      6285  CACTGTGTAAAAATTTAGCAATACTACATACACAACCACTCCCAATGCATCAGGTGGGT 6344
|||||
Db      457  CACTGTGTAAAAATTTAGCAATACTACATACACAACCACTCCCAATGCATCAGGTGGGT 398
|||||
QY      6345  AACTCTCCCAACAATAAGTCTGCCTACCCCTCAGGAATATTTTGTCTGTGGTACCTC 6404
|||||
Db      397  AACTCTCCCAACAATAAGTCTGCCTACCCCTCAGGAATATTTTGTCTGTGGTACCTC 338
|||||
QY      6405  AGCCTATCGTTTGTGAATGGCTCTTCAGAAATCTATGTGCTTCCTCTCAATCTTAGTGC 6464
|||||
Db      337  AGCCTATCGTTTGTGAATGGCTCTTCAGAAATCTATGTGCTTCCTCTCAATCTTAGTGC 278
|||||

QY      6465  CCCYATGRCATCTACACTGAACAAGATTATATACAGTTATGTATCTAATCTAAGCCCGCAA 6524
|||||
Db      277  CCCTATGACCATCTACACTGAACAAGATTATATACAAATTATGTATATCTAAGCCCGCAA 218
|||||
QY      6525  CAAAAGAGTACCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6584
|||||
Db      217  CAAAAGAGTACCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 158
|||||
QY      6585  TGGCATTGGCGGTATCACAACTCTACTAGTTCTACTACAAAATCTCTCAAGAATACTAAA 6644
|||||
Db      157  TGGCATTGGCGGTATCACAACTCTACTAGTTCTACTACAAAATCTCTCAAGAATACTAAA 98
|||||
QY      6645  TGGGGACATGGAACGGGTGCGGACTCCCTGGTGCACCTTGCACAAAGATCAACTTAATCCCT 6704
|||||
Db      97  TGGGGACATGGAACGGGTGCGGACTCCCTGGTGCACCTTGCACAAAGATCAACTTAATCCCT 38
|||||
QY      6705  AGCAGCAGTAGTCTCTTC 6721
|||||
Db      37  AGCAGCAGTAGTCTCTTC 21
|||||

RESULT 28
BX347952
LOCUS      BX347952
DEFINITION      955 bp      mRNA      linear      EST 08-APR-2004
clone CS0DI051YM13 5-PRIME, mRNA sequence.
ACCESSION      BX347952
VERSION      BX347952.1 GI:30375235
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 955)
Li, W. B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAB0232F07_AE02167_l&c=4215.r

FEATURES
source      Location/Qualifiers
1. .955
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI051YM13"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoRV
sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN
Query Match      4.5%; Score 344; DB 4; Length 955;
Best Local Similarity 99.3%; Pred. No. 3.9e-180;
Matches 544; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      5806  GGCACCTTTTGGATGCAATCAATCTATTTAGCAGGAAAAATGATTAATCTTAGT 5865
|||||
Db      1  GGCACCTTTTGGATGCAATCAATCTATTTAGCAGGAAAAATGATTAATCTTAGT 60
|||||

```

```
QY 5866 TGTCTCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTACCCAACTGGTATGCT 5925
Db |||||
61 TGTCTCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTACCCAACTGGTATGCT 120
QY 5926 GATGGGGTGGAGTTCAGAGTCAGGCAGAGAGAAAAACATGTAAAAAGAGTATCTCCCAA 5985
Db |||||
121 GATGGGGTGGAGTTCAGAGTCAGGCAGAGAGAAAAACATGTAAAAAGAGTATCTCCCAA 180
QY 5986 CTCACSSGGGTACATGGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACAT 6045
Db |||||
181 CTCACSSGGGTACATGGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACAT 240
QY 6046 GAAACCTCTCCGTACCCACTACTCGCTGTAGCCCTTAAATACACACCTCTAGGGTCT 6105
Db |||||
241 GAAACCTCTCCGTACCCACTACTCGCTGTAGCCCTTAAATACACACCTCTAGGGTCT 300
QY 6106 CATGAGGTCTCGGCCCAAAACCTTACTACTGTGTGGATGCTCCCTCGAACTTCARG 6165
Db |||||
301 CATGAGGTCTCGGCCCAAAACCTTACTACTGTGTGGATGCTCCCTCGAACTTCARG 360
QY 6166 CCATATGTTTCAATCCCTGTACTGAAACAATGGAACAACCTTCAGCACAGAAATAAACACC 6225
Db |||||
361 CCATATGTTTCAATCCCTGTACTGAAACAATGGAACAACCTTCAGCACAGAAATAAACACC 420
QY 6226 ACTTCGGTTTATAGTAGGACCTCTGTTTCCAAATSTGGAAATTAACCATACCTCAAACTCT 6285
Db |||||
421 ACTTCGGTTTATAGTAGGACCTCTGTTTCCAAATSTGGAAATTAACCATACCTCAAACTCT 480
QY 6286 ACCTGTGTAATTTAGCAATACTACATACACAACCACTCCCAATGATCAGTGGGTA 6345
Db |||||
481 ACCTGTGTAATTTAGCAATACTACATACACAACCACTCCCAATGATCAGTGGGTA 540
QY 6346 ACTCCTCC 6353
Db |||||
541 ACTCCTCC 548
```

```
RESULT 29
CR735592
LOCUS
DEFINITION
CR735592 NIH_MGC 21 Homo sapiens cDNA clone IMAGE958M21530 ;
IMAGE:3840572 5', mRNA sequence.
ACCESSION
CR735592
VERSION
CR735592.1 GI:51584854
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 554)
Ebert,L., Heil,O., Hennig,S., Korn,B., Neubert,P., Partsch,E.,
Peters,M., Radelof,U. and Schneider,D.
I.M.A.G.E. cDNA Clone Collection
Unpublished (2004)
Contact: Inge Ariart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD; IMAGE958M21530.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Contact: Inge Ariart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTCACAGGAAACAGCTATGAC.
FEATURES
Location/Qualifiers
1..554
```

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="IMAGE958M21530 ; IMAGE:3840572"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 21"
/notes="Organ: placenta; Vector: pOTB7; Site:1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

	Query Match	4.5%	Score 338;	DB 8;	Length 554;
	Best Local Similarity	99.5%;	Pred. No. 8.8e-177;		
	Matches 438;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	6282	CTCACCTGTGTAAAAATTTAGCAATATACATACACAACCAACTCCCAATGCATCAGGTG	6341		
Db	38	CTCACCTGTGTAAAAATTTAGCAATATACATACACAACCAACTCCCAATGCATCAGGTG	97		
QY	6342	GGTAATCTCTCCACACAAATAGTCTGCTACCTCCAGCAATATTTTGTCTGTGGTAC	6401		
Db	98	GGTAATCTCTCCACACAAATAGTCTGCTACCTCCAGCAATATTTTGTCTGTGGTAC	157		
QY	6402	CTCAGCCTATCGTGTGTTTGAATGGCTCTTCAGAATCTATGTCTCTCTCATCTTAGT	6461		
Db	158	CTCAGCCTATCGTGTGTTTGAATGGCTCTTCAGAATCTATGTCTCTCTCATCTTAGT	217		
QY	6462	GCCCCYATGCCATCTACACTGAAACAAGATTTATACAGTTATGTCATATCTAAGCCCG	6521		
Db	218	GCCCCYATGCCATCTACACTGAAACAAGATTTATACAGTTATGTCATATCTAAGCCCG	277		
QY	6522	CAACAAAGAGTACCATTCTTCTTTTATAGGACGAGGCTAGGTGCACCTAGG	6581		
Db	278	CAACAAAGAGTACCATTCTTCTTTTATAGGACGAGGCTAGGTGCACCTAGG	337		
QY	6582	TACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTTACTACAAACTATCTCAAGAACT	6641		
Db	338	TACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTTACTACAAACTATCTCAAGAACT	397		
QY	6642	AAATGGGACATGGAAACGGGTGCGGACTCCCTGCTCAGTTGCAAGATCAACTTAATCTC	6701		
Db	398	AAATGGGACATGGAAACGGGTGCGGACTCCCTGCTCAGTTGCAAGATCAACTTAATCTC	457		
QY	6702	CCTAGCAGCAGTAGTCCCTTC	6721		
Db	458	CCTAGCAGCAGTAGTCCCTTC	477		
RESULT 30					
LOCUS	924 bp	mRNA	linear	EST 01-MAY-2004	
DEFINITION	BX409328 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE013120				
ACCESSION	5-PRIME, mRNA sequence.				
VERSION	BX409328				
KEYWORDS	EST.				
SOURCE	EST.				
ORGANISM	Homo sapiens (human)				
REFERENCE	1 (bases 1 to 924)				
AUTHORS	Li,W.B., Gruber,C., Jessee,J. and Polayes,D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished (2001)				
COMMENT	On May 13, 2003 this sequence version replaced gi:30652931.				
	Contact: Genoscope				



Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
was not normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
This sequence belongs to sequence cluster 4215.r  
For more information about this cluster, see  
[http://www.genoscope.cns.fr/cdna?s=CS0BAE0122F02\\_AE01071\\_l&c=4215.r](http://www.genoscope.cns.fr/cdna?s=CS0BAE0122F02_AE01071_l&c=4215.r)

## FEATURES

source	Location/Qualifiers
1. .924	
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="CS0DB013Y120"
	/tissue_type="PLACENTA"
	/clone_lib="Homo sapiens PLACENTA"
	/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

## ORIGIN

Query Match	4.4%; Score 336; DB 4; Length 924;
Best Local Similarity	99.5%; Pred. No. 1.2e-175;
Matches 436; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY 5685	AGAGTTTCTATGAGAAATGAGCGTCCCGAAATATTGATGCCCATCGTATAGGAGTCT 5744
Db 155	AGAGTTTCTATGAGAAATGAGCGTCCCGAAATATTGATGCCCATCGTATAGGAGTCT 214
QY 5745	TTSTAAGGGAACCCACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 5804
Db 215	TTCTAAGGGAACCCACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 274
QY 5805	TGCCACTCTTTGTCATGCATGCAAAATCTCATTTATGGACAGGAAATGATTAATCCTAG 5864
Db 275	TGCCACTCTTTGTCATGCATGCAAAATCTCATTTATGGACAGGAAATGATTAATCCTAG 334
QY 5865	TTGTCTTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCAACTGGTAGTC 5924
Db 335	TTGTCTTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCAACTGGTAGTC 394
QY 5925	TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAGAAGTAAATCTCCCA 5984
Db 395	TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAGAAGTAAATCTCCCA 454
QY 5985	ACTCACCGGGTACATGGGACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACA 6044
Db 455	ACTCACCGGGTACATGGGACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACA 514
QY 6045	TGAACCTCCGTACCCATCTCGCTGGTAAAGCTTATTAAATACCACTCTCACTGGGCT 6104
Db 515	TGAACCTCCGTACCCATCTCGCTGGTAAAGCTTATTAAATACCACTCTCACTGGGCT 574
QY 6105	CCATGAGGTCTCGGCCCA 6122
Db 575	CCATGAGGTCTCGGCCCA 592

Search completed: August 2, 2006, 15:18:55  
Job time : 21283 secs

**This Page Blank (uspto)**

GenCore version 5.1.9  
 Copyright (c) 1993 - 2006 Bioceleration Ltd.  
 OM nucleic - nucleic search, using sw model  
 Run on: August 2, 2006, 09:10:45 ; Search time 436 Seconds  
 (without alignments)  
 9710.509 Million cell updates/sec

Title: US-10-717-580-11  
 Perfect score: 7582  
 Sequence: 1 caacaatcggtataaacc.....tattaatcttcgactgtg 7582

Scoring table: OLIGO\_NUC  
 Gapop 60.0 , Gapext 60.0  
 Searched: 1452992 seqs, 279199553 residues

Word size : 6  
 Total number of hits satisfying chosen parameters: 2903664  
 Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Listing first 90 summaries

Database : Published Applications NA New.\*  
 1: /EMC\_Celerra\_SID33/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
 2: /EMC\_Celerra\_SID33/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
 3: /EMC\_Celerra\_SID33/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
 4: /EMC\_Celerra\_SID33/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
 5: /EMC\_Celerra\_SID33/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
 6: /EMC\_Celerra\_SID33/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
 7: /EMC\_Celerra\_SID33/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
 8: /EMC\_Celerra\_SID33/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	0.4	44063	6	US-10-539-228-718
2	28	0.4	138627	6	US-10-540-898-159
3	22	0.3	35	6	US-10-544-059-33
4	22	0.3	36	6	US-10-544-059-41
5	22	0.3	39	6	US-10-544-059-30
6	21	0.3	36	6	US-10-544-059-37
7	21	0.3	40	6	US-10-544-059-32
8	21	0.3	170489	6	US-10-539-228-380
9	20	0.3	20	6	US-10-554-711-771
10	20	0.3	20	6	US-10-554-711-775
11	20	0.3	38	6	US-10-544-059-36
12	20	0.3	50	6	US-10-554-711-192
13	20	0.3	1507	7	US-11-056-3558-55306
14	20	0.3	256190	6	US-10-539-228-320
15	19	0.3	35	6	US-10-544-059-40
16	19	0.3	35	6	US-10-544-059-42
17	19	0.3	35	6	US-10-544-059-44
18	19	0.3	1877	6	US-10-449-902-25572
19	19	0.3	2047	6	US-10-516-094-14
20	19	0.3	2994	6	US-10-540-898-5
21	19	0.3	84476	6	US-10-540-898-4
22	19	0.3	225587	6	US-10-540-898-374
23	18	0.2	18	6	US-10-554-711-777
24	18	0.2	33	6	US-10-544-059-31
25	18	0.2	33	6	US-10-544-059-39

Sequence 43, Appl	33	0.2	18	26	US-10-544-059-43	6	Sequence 43, Appl
Sequence 34, Appl	40	0.2	18	27	US-10-544-059-34	6	Sequence 34, Appl
Sequence 23, Appl	561	0.2	18	28	US-11-300-928-23	7	Sequence 23, Appl
Sequence 2900, Ap	575	0.2	18	29	US-10-488-619-2900	7	Sequence 2900, Ap
Sequence 2441, A	840	0.2	18	30	US-11-218-305-2441	7	Sequence 2441, A
Sequence 4897, Ap	1080	0.2	18	31	US-10-471-571A-4897	6	Sequence 4897, Ap
Sequence 557, App	1095	0.2	18	32	US-10-471-571A-557	6	Sequence 557, App
Sequence 58368, A	1099	0.2	18	33	US-11-056-3558-58368	7	Sequence 58368, A
Sequence 2644, Ap	1109	0.2	18	34	US-10-449-902-2644	6	Sequence 2644, Ap
Sequence 2912, Ap	1139	0.2	18	35	US-11-218-305-2912	7	Sequence 2912, Ap
Sequence 2033, Ap	1223	0.2	18	36	US-10-449-902-2033	6	Sequence 2033, Ap
Sequence 13693, A	1273	0.2	18	37	US-11-218-305-13693	7	Sequence 13693, A
Sequence 26581, A	1385	0.2	18	38	US-10-449-902-26581	6	Sequence 26581, A
Sequence 2315, Ap	1382	0.2	18	39	US-10-449-902-2315	6	Sequence 2315, Ap
Sequence 21352, A	1377	0.2	18	40	US-10-449-902-21352	6	Sequence 21352, A
Sequence 1867, Ap	1395	0.2	18	41	US-10-449-902-1867	6	Sequence 1867, Ap
Sequence 23075, A	1420	0.2	18	42	US-10-449-902-23075	6	Sequence 23075, A
Sequence 947, App	1426	0.2	18	43	US-10-449-902-947	6	Sequence 947, App
Sequence 3077, Ap	1463	0.2	18	44	US-10-449-902-3077	6	Sequence 3077, Ap
Sequence 16130, A	1480	0.2	18	45	US-10-449-902-16130	6	Sequence 16130, A
Sequence 13694, A	1511	0.2	18	46	US-11-218-305-13694	7	Sequence 13694, A
Sequence 17084, A	1559	0.2	18	47	US-11-218-305-17084	7	Sequence 17084, A
Sequence 64677, A	1572	0.2	18	48	US-11-056-3558-64677	7	Sequence 64677, A
Sequence 1162, Ap	1634	0.2	18	49	US-10-449-902-1162	6	Sequence 1162, Ap
Sequence 4631, Ap	1664	0.2	18	50	US-10-449-902-4631	6	Sequence 4631, Ap
Sequence 26973, A	1684	0.2	18	51	US-10-449-902-26973	6	Sequence 26973, A
Sequence 2365, Ap	1779	0.2	18	52	US-10-449-902-2365	6	Sequence 2365, Ap
Sequence 1149, Ap	1822	0.2	18	53	US-10-449-902-1149	6	Sequence 1149, Ap
Sequence 17925, A	1872	0.2	18	54	US-10-449-902-17925	6	Sequence 17925, A
Sequence 14299, A	1896	0.2	18	55	US-10-449-902-14299	6	Sequence 14299, A
Sequence 28272, A	2132	0.2	18	56	US-10-449-902-28272	6	Sequence 28272, A
Sequence 29486, A	2162	0.2	18	57	US-10-449-902-29486	6	Sequence 29486, A
Sequence 33076, A	2187	0.2	18	58	US-10-449-902-33076	6	Sequence 33076, A
Sequence 8396, Ap	2369	0.2	18	59	US-11-056-3558-8396	7	Sequence 8396, Ap
Sequence 87454, A	2390	0.2	18	60	US-10-449-902-87454	7	Sequence 87454, A
Sequence 9934, Ap	2448	0.2	18	61	US-11-218-305-9934	7	Sequence 9934, Ap
Sequence 6620, Ap	2527	0.2	18	62	US-10-449-902-6620	6	Sequence 6620, Ap
Sequence 269, App	2589	0.2	18	63	US-10-449-902-269	6	Sequence 269, App
Sequence 23130, A	2597	0.2	18	64	US-10-539-228-23130	6	Sequence 23130, A
Sequence 13810, A	2905	0.2	18	65	US-10-449-902-13810	6	Sequence 13810, A
Sequence 12645, A	3145	0.2	18	66	US-11-218-305-9936	7	Sequence 12645, A
Sequence 18941, A	3614	0.2	18	67	US-10-449-902-18941	6	Sequence 18941, A
Sequence 744, App	4025	0.2	18	68	US-10-449-902-744	6	Sequence 744, App
Sequence 596, App	7230	0.2	18	69	US-10-540-898-596	6	Sequence 596, App
Sequence 137, App	32847	0.2	18	70	US-10-505-928-137	6	Sequence 137, App
Sequence 268, App	70665	0.2	18	71	US-10-540-898-268	6	Sequence 268, App
Sequence 835, App	107432	0.2	18	72	US-10-539-228-835	6	Sequence 835, App
Sequence 729, App	107745	0.2	18	73	US-10-540-898-729	6	Sequence 729, App
Sequence 1, Appl	180057	0.2	18	74	US-10-540-898-1	6	Sequence 1, Appl
Sequence 1043443	180026	0.2	18	75	US-10-540-898-1043443	6	Sequence 1043443
Sequence 10, Appl	195102	0.2	18	76	US-10-733-042-1	6	Sequence 10, Appl
Sequence 126, App	257645	0.2	18	77	US-10-540-898-126	6	Sequence 126, App
Sequence 12265, A	25	0.2	17	78	US-11-348-413-12265	7	Sequence 12265, A
Sequence 29, Appl	50	0.2	17	79	US-10-511-937-29	6	Sequence 29, Appl
Sequence 78211, A	240	0.2	17	80	US-11-348-711-78211	7	Sequence 78211, A
Sequence 1539, Ap	494	0.2	17	81	US-11-376-673-1539	7	Sequence 1539, Ap
Sequence 66, Appl	501	0.2	17	82	US-10-488-619-66	6	Sequence 66, Appl
Sequence 69, Appl	574	0.2	17	83	US-10-541-993-69	6	Sequence 69, Appl
Sequence 7472, Ap	611	0.2	17	84	US-10-541-993-7472	6	Sequence 7472, Ap
	723	0.2	17	85	US-10-449-902-723	6	

ALIGNMENTS

RESULT 1  
 US-10-539-228-718  
 ; Sequence 718, Application us/10539228

; Publication No. US20060154250A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc S. Malandro  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
; FILE REFERENCE: CHIR0052-101 (PP023370.0003)  
; CURRENT APPLICATION NUMBER: US/10/539,228  
; CURRENT FILING DATE: 2005-06-17  
; PRIOR APPLICATION NUMBER: US 10/322,281  
; PRIOR FILING DATE: 2002-12-18  
; NUMBER OF SEQ ID NOS: 866  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 718  
; LENGTH: 44063  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(44063)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-539-228-718

Query Match 0.4%; Score 34; DB 6; Length 44063;  
Best Local Similarity 100.0%; Pred. No. 1.1e-07;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5709 TCCCGGAATATGTATGCCCATCGTATAGGAGT 5742  
|||||  
DB 916 TCCCGGAATATGTATGCCCATCGTATAGGAGT 949  
|||||

## RESULT 2

US-10-540-898-159/c  
; Sequence 159, Application US/10540898  
; Publication No. US20060166213A1

; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc Malandro  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
; FILE REFERENCE: CHIR0056-101 (PP023367.0003)  
; CURRENT APPLICATION NUMBER: US/10/540,898  
; CURRENT FILING DATE: 2005-06-27  
; PRIOR APPLICATION NUMBER: US 10/330,773  
; PRIOR FILING DATE: 2002-12-27  
; NUMBER OF SEQ ID NOS: 981  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 159  
; LENGTH: 138627  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(138627)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-540-898-159

Query Match 0.4%; Score 28; DB 6; Length 138627;  
Best Local Similarity 100.0%; Pred. No. 0.00025;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 GGGATATAAACCCAGGCATTCGAGCTGG 36  
|||||  
DB 81809 GGGATATAAACCCAGGCATTCGAGCTGG 81782  
|||||

## RESULT 3

US-10-544-059-33  
; Sequence 33, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; APPLICANT: Hadlock, Kenneth G.

; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; TITLE OF INVENTION: LATERAL SCLEROSIS  
; FILE REFERENCE: 376462002040  
; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 33  
; LENGTH: 35  
; TYPE: DNA  
; ORGANISM: Herv-W  
US-10-544-059-33

Query Match 0.3%; Score 22; DB 6; Length 35;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6614 AGTTCTACTACAAACTATCTCA 6635  
|||||  
DB 14 AGTTCTACTACAAACTATCTCA 35  
|||||

## RESULT 4

US-10-544-059-41  
; Sequence 41, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; TITLE OF INVENTION: LATERAL SCLEROSIS  
; FILE REFERENCE: 376462002040  
; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 41  
; LENGTH: 36  
; TYPE: DNA  
; ORGANISM: Herv-W  
US-10-544-059-41

Query Match 0.3%; Score 22; DB 6; Length 36;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1270 GTTTGGCGATCTCTGGTATCTC 1291  
|||||  
DB 15 GTTTGGCGATCTCTGGTATCTC 36  
|||||

## RESULT 5

US-10-544-059-30/c  
; Sequence 30, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; TITLE OF INVENTION: LATERAL SCLEROSIS  
; FILE REFERENCE: 376462002040  
; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 30  
; LENGTH: 39  
; TYPE: DNA  
; ORGANISM: Herv-W  
US-10-544-059-30

Query Match 0.3%; Score 22; DB 6; Length 39;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6135 CTCTGGATATGCTCCCTG 6156  
|||||  
Db 39 CTCTGGATATGCTCCCTG 18

RESULT 6  
US-10-544-059-37  
; Sequence 37, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; FILE REFERENCE: LATERAL SCLEROSIS  
; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 37  
; LENGTH: 36  
; TYPE: DNA  
; ORGANISM: Herv-W  
US-10-544-059-37

Query Match 0.3%; Score 21; DB 6; Length 36;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 963 CAAAGGAGATACACAAAGG 983  
|||||  
Db 16 CAAAGGAGATACACAAAGG 36

RESULT 7  
US-10-544-059-32/C  
; Sequence 32, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; FILE REFERENCE: LATERAL SCLEROSIS  
; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 32  
; LENGTH: 40  
; TYPE: DNA  
; ORGANISM: Herv-W  
US-10-544-059-32

Query Match 0.3%; Score 21; DB 6; Length 40;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6512 CTAAGCCCGCAACAAAGAG 6532  
|||||  
Db 40 CTAAGCCCGCAACAAAGAG 20

RESULT 8  
US-10-539-228-380  
; Sequence 380, Application US/10539228  
; Publication No. US20060154250A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc S. Malandro  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
; FILE REFERENCE: CHIR0052-101 (PP023370.0003)  
; CURRENT APPLICATION NUMBER: US/10/539,228  
; CURRENT FILING DATE: 2005-06-17  
; PRIOR APPLICATION NUMBER: US 10/322,281  
; PRIOR FILING DATE: 2002-12-18  
; NUMBER OF SEQ ID NOS: 866  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 380  
; LENGTH: 170489  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)..(170489)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-539-228-380

Query Match 0.3%; Score 21; DB 6; Length 170489;  
Best Local Similarity 100.0%; Pred. No. 1.9;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5994 GGTACATGGCACCTCTAGCCC 6014  
|||||  
Db 167429 GGTACATGGCACCTCTAGCCC 167449

RESULT 9  
US-10-554-711-771  
; Sequence 771, Application US/10554711  
; Publication No. US20060115806A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Georgia Research Foundation  
; APPLICANT: McDonald, John F.  
; TITLE OF INVENTION: Global Analysis of Transposable Elements  
; FILE REFERENCE: as Molecular Markers of Cancer  
; CURRENT APPLICATION NUMBER: US/10/554,711  
; CURRENT FILING DATE: 2005-10-28  
; PRIOR APPLICATION NUMBER: 60/466,798  
; PRIOR FILING DATE: 2003-04-29  
; NUMBER OF SEQ ID NOS: 778  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 771  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:/note =  
; OTHER INFORMATION: Synthetic Construct  
US-10-554-711-771

Query Match 0.3%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 171 CCACCACCTGCTTTGCCAC 190  
|||||  
Db 1 CCACCACCTGCTTTGCCAC 20

```
RESULT 10
US-10-554-711-775
; Sequence 775, Application US/10554711
; Publication No. US20060115806A1
; GENERAL INFORMATION:
; APPLICANT: University of Georgia Research Foundation
; APPLICANT: McDonald, John F.
; TITLE OF INVENTION: Global Analysis of Transposable Elements
; TITLE OF INVENTION: as Molecular Markers of Cancer
; FILE REFERENCE: 21099.0075P1
; CURRENT APPLICATION NUMBER: US/10/554,711
; CURRENT FILING DATE: 2005-10-28
; PRIOR APPLICATION NUMBER: 60/466,798
; PRIOR FILING DATE: 2003-04-29
; NUMBER OF SEQ ID NOS: 778
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 775
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence/note =
; OTHER INFORMATION: Synthetic Construct
US-10-554-711-775

Query Match      0.3%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6590 TTGGCGGTATCATCAACCTCT 6609
Db 1 TTGGCGGTATCATCAACCTCT 20

RESULT 11
US-10-544-059-36/c
; Sequence 36, Application US/10544059
; Publication No. US20060160087A1
; GENERAL INFORMATION:
; APPLICANT: SLIL Biomedical Corporation
; APPLICANT: McGrath, Michael
; APPLICANT: Hadlock, Kenneth G.
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC
; TITLE OF INVENTION: LATERAL SCLEROSIS
; FILE REFERENCE: 376462002040
; CURRENT APPLICATION NUMBER: US/10/544,059
; CURRENT FILING DATE: 2005-07-29
; PRIOR APPLICATION NUMBER: US 60/444,224
; PRIOR FILING DATE: 2003-01-31
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 38
; TYPE: DNA
; ORGANISM: Herv-W
US-10-544-059-36

Query Match      0.3%; Score 20; DB 6; Length 38;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1005 GCCAATATTCCTCCCAATTATG 1024
Db 38 GCCAATATTCCTCCCAATTATG 19

RESULT 12
US-10-554-711-192
; Sequence 192, Application US/10554711
; Publication No. US20060115806A1
; GENERAL INFORMATION:
; APPLICANT: University of Georgia Research Foundation
```

```
; APPLICANT: McDonald, John F.
; TITLE OF INVENTION: Global Analysis of Transposable Elements
; TITLE OF INVENTION: as Molecular Markers of Cancer
; FILE REFERENCE: 21099.0075P1
; CURRENT APPLICATION NUMBER: US/10/554,711
; CURRENT FILING DATE: 2005-10-28
; PRIOR APPLICATION NUMBER: 60/466,798
; PRIOR FILING DATE: 2003-04-29
; NUMBER OF SEQ ID NOS: 778
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 192
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence/note =
; OTHER INFORMATION: Synthetic Construct
US-10-554-711-192

Query Match      0.3%; Score 20; DB 6; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 484 GCTTGCCACCATCTTGGGAAG 503
Db 29 GCTTGCCACCATCTTGGGAAG 48

RESULT 13
US-11-056-355B-55306
; Sequence 55306, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590FUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 55306
; LENGTH: 1507
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1507)
; OTHER INFORMATION: Ceres Seq. ID no. 12607536
US-11-056-355B-55306

Query Match      0.3%; Score 20; DB 7; Length 1507;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6439 ATGTGTTTCCTCTCATCTT 6458
Db 736 ATGTGTTTCCTCTCATCTT 755

RESULT 14
US-10-539-228-320/c
; Sequence 320, Application US/10539228
; Publication No. US20060154250A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: CHIR0052-101 (PP023370.0003)
; CURRENT APPLICATION NUMBER: US/10/539,228
; CURRENT FILING DATE: 2005-06-17
```



; PRIOR APPLICATION NUMBER: US 10/322,281  
; PRIOR FILING DATE: 2002-12-18  
; NUMBER OF SEQ ID NOS: 866  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 320  
; LENGTH: 256190  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(256190)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-539-228-320

Query Match 0.3%; Score 20; DB 6; Length 256190;  
Best Local Similarity 100.0%; Pred. No. 6.9;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1502 CACAGGGGAAGGAAGAAA 1521  
|||||  
DB 88993 CACAGGGGAAGGAAGAAA 88974

## RESULT 15

US-10-544-059-40/c  
; Sequence 40, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; FILE REFERENCE: 376462002040  
; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 40  
; LENGTH: 35  
; TYPE: DNA  
; ORGANISM: Herv-W  
US-10-544-059-40

Query Match 0.3%; Score 19; DB 6; Length 35;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1156 GTTTTCAAGGGTTAGGAC 1174  
|||||  
DB 35 GTTTTCAAGGGTTAGGAC 17

## RESULT 16

US-10-544-059-42/c  
; Sequence 42, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; FILE REFERENCE: 376462002040  
; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 42

; LENGTH: 35  
; TYPE: DNA  
; ORGANISM: Herv-W  
US-10-544-059-42

Query Match 0.3%; Score 19; DB 6; Length 35;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1520 AATCCTACTGCCCTTCTGG 1538  
|||||  
DB 35 AATCCTACTGCCCTTCTGG 17

## RESULT 17

US-10-544-059-44/c  
; Sequence 44, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; FILE REFERENCE: 376462002040  
; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 44  
; LENGTH: 35  
; TYPE: DNA  
; ORGANISM: Herv-W  
US-10-544-059-44

Query Match 0.3%; Score 19; DB 6; Length 35;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1622 ATCACTCAGTCAGCTGCAG 1640  
|||||  
DB 35 ATCACTCAGTCAGCTGCAG 17

## RESULT 18

US-10-449-902-25572  
; Sequence 25572, Application US/10449902  
; Publication No. US20060123505A1  
; GENERAL INFORMATION:  
; APPLICANT: National Institute of Agrobiological Sciences.  
; APPLICANT: Bio-oriented Technology Research Advancement Institution.  
; APPLICANT: The Institute of Physical and Chemical Research.  
; APPLICANT: Foundation for Advancement of International Science.  
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF  
; FILE REFERENCE: MOA-A0205V1-US  
; CURRENT APPLICATION NUMBER: US/10/449,902  
; CURRENT FILING DATE: 2003-05-29  
; PRIOR APPLICATION NUMBER: JP 2002-203269  
; PRIOR FILING DATE: 2002-05-30  
; PRIOR APPLICATION NUMBER: JP 2002-383870  
; PRIOR FILING DATE: 2002-12-11  
; NUMBER OF SEQ ID NOS: 56791  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 25572  
; LENGTH: 1877  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: AK101014  
; DATABASE ENTRY DATE: 2002-08-28  
US-10-449-902-25572

Query Match 0.3%; Score 19; DB 6; Length 1877;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1533 TTCTGGAGAGACTAAGGGA 1551

DB 419 TTCTGGAGAGACTAAGGGA 437

## RESULT 19

US-10-516-094-14/c  
; Sequence 14, Application US/10516094  
; Publication No. US20060141457A1  
; GENERAL INFORMATION:  
; APPLICANT: Dahlqvist, Anders  
; APPLICANT: Ghosal, Alakesh  
; APPLICANT: Lindqvist, Ylva  
; APPLICANT: Banas, Antoni  
; TITLE OF INVENTION: NEW IMPROVED ACYLTRANSFERASE  
; FILE REFERENCE: 5822.317USWO  
; CURRENT APPLICATION NUMBER: US/10/516,094  
; CURRENT FILING DATE: 2004-11-24  
; PRIOR APPLICATION NUMBER: PCT/SE03/00870  
; PRIOR FILING DATE: 2003-05-28  
; PRIOR APPLICATION NUMBER: SE 0201581-6  
; PRIOR FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: SE 0300142-7  
; PRIOR FILING DATE: 2003-01-20  
; PRIOR APPLICATION NUMBER: US 60/383,889  
; PRIOR FILING DATE: 2002-05-29  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 14  
; LENGTH: 2047  
; TYPE: DNA  
; ORGANISM: Aspergillus nidulans  
US-10-516-094-14

Query Match 0.3%; Score 19; DB 6; Length 2047;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6442 TGCTTCCTCTCATCTTAG 6460

DB 396 TGCTTCCTCTCATCTTAG 378

## RESULT 20

US-10-540-898-5/c  
; Sequence 5, Application US/10540898  
; Publication No. US20060166213A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc Malandro  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
; FILE REFERENCE: CHIR0056-101 (PP023367.0003)  
; CURRENT APPLICATION NUMBER: US/10/540,898  
; CURRENT FILING DATE: 2005-06-27  
; PRIOR APPLICATION NUMBER: US 10/330,773  
; PRIOR FILING DATE: 2002-12-27  
; NUMBER OF SEQ ID NOS: 981  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 2994  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-540-898-5

Query Match 0.3%; Score 19; DB 6; Length 2994;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1976 TGCCCCAGGGGACAAAGGT 1994

DB 1594 TGCCCCAGGGGACAAAGGT 1576

## RESULT 21

US-10-540-898-4/c  
; Sequence 4, Application US/10540898  
; Publication No. US20060166213A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc Malandro  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
; FILE REFERENCE: CHIR0056-101 (PP023367.0003)  
; CURRENT APPLICATION NUMBER: US/10/540,898  
; CURRENT FILING DATE: 2005-06-27  
; PRIOR APPLICATION NUMBER: US 10/330,773  
; PRIOR FILING DATE: 2002-12-27  
; NUMBER OF SEQ ID NOS: 981  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 84476  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-540-898-4

Query Match 0.3%; Score 19; DB 6; Length 84476;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1976 TGCCCCAGGGGACAAAGGT 1994

DB 73076 TGCCCCAGGGGACAAAGGT 73058

## RESULT 22

US-10-540-898-374/c  
; Sequence 374, Application US/10540898  
; Publication No. US20060166213A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc Malandro  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
; FILE REFERENCE: CHIR0056-101 (PP023367.0003)  
; CURRENT APPLICATION NUMBER: US/10/540,898  
; CURRENT FILING DATE: 2005-06-27  
; PRIOR APPLICATION NUMBER: US 10/330,773  
; PRIOR FILING DATE: 2002-12-27  
; NUMBER OF SEQ ID NOS: 981  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 374  
; LENGTH: 225587  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(225587)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-540-898-374

Query Match 0.3%; Score 19; DB 6; Length 225587;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6992 TCCTTGTAACTTGTCTC 7010

DB 218645 TCCTTGTAACTTGTCTC 218627

## RESULT 23

US-10-554-711-777/c  
; Sequence 777, Application US/10554711  
; Publication No. US20060115806A1

; GENERAL INFORMATION:  
; APPLICANT: University of Georgia Research Foundation  
; APPLICANT: McDonald, John F.  
; TITLE OF INVENTION: Global Analysis of Transposable Elements  
; FILE REFERENCE: 21099.0075P1  
; CURRENT APPLICATION NUMBER: US/10/554,711  
; PRIOR FILING DATE: 2005-10-28  
; PRIOR APPLICATION NUMBER: 60/466,798  
; PRIOR FILING DATE: 2003-04-29  
; NUMBER OF SEQ ID NOS: 778  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 777  
; LENGTH: 18  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:/note =  
; OTHER INFORMATION: Synthetic Construct  
US-10-554-711-777

Query Match 0.2%; Score 18; DB 6; Length 18;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6807 TCAATCCGGAATCGTCAC 6824  
Db 18 TCAATCCGGAATCGTCAC 1

RESULT 24  
US-10-544-059-31  
; Sequence 31, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; TITLE OF INVENTION: LATERAL SCLEROSIS  
; FILE REFERENCE: 376462002040  
; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 31  
; LENGTH: 33  
; TYPE: DNA  
; ORGANISM: Herv-W  
US-10-544-059-31

Query Match 0.2%; Score 18; DB 6; Length 33;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5935 GGAGTTCAAGATCAGGCA 5952  
Db 16 GGAGTTCAAGATCAGGCA 33

RESULT 25  
US-10-544-059-39  
; Sequence 39, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; TITLE OF INVENTION: LATERAL SCLEROSIS  
; FILE REFERENCE: 376462002040

; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 39  
; LENGTH: 33  
; TYPE: DNA  
; ORGANISM: Herv-W  
US-10-544-059-39

Query Match 0.2%; Score 18; DB 6; Length 33;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 866 TTATGCCCTACAGGAAGC 883  
Db 16 TTATGCCCTACAGGAAGC 33

RESULT 26  
US-10-544-059-43  
; Sequence 43, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; TITLE OF INVENTION: LATERAL SCLEROSIS  
; FILE REFERENCE: 376462002040  
; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 43  
; LENGTH: 33  
; TYPE: DNA  
; ORGANISM: Herv-W  
US-10-544-059-43

Query Match 0.2%; Score 18; DB 6; Length 33;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1486 ATGATGTCCACCAATACA 1503  
Db 16 ATGATGTCCACCAATACA 33

RESULT 27  
US-10-544-059-34/c  
; Sequence 34, Application US/10544059  
; Publication No. US20060160087A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIL Biomedical Corporation  
; APPLICANT: McGrath, Michael  
; APPLICANT: Hadlock, Kenneth G.  
; TITLE OF INVENTION: MONITORING AND TREATMENT OF AMYOTROPHIC  
; TITLE OF INVENTION: LATERAL SCLEROSIS  
; FILE REFERENCE: 376462002040  
; CURRENT APPLICATION NUMBER: US/10/544,059  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US 60/444,224  
; PRIOR FILING DATE: 2003-01-31  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 34  
; LENGTH: 40  
; TYPE: DNA

```

; ORGANISM: Herv-W
US-10-544-059-34

Query Match          0.2%; Score 18; DB 6; Length 40;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6890 CCTGGGGCCCTCTCAGCC 6907
Db 40 CCTGGGGCCCTCTCAGCC 23

RESULT 28
US-11-300-928-23
; Sequence 23, Application US/11300928
; Publication No. US20060166277A1
; GENERAL INFORMATION:
; APPLICANT: Karumanchi, S. Ananth
; APPLICANT: Sukhatme, Vikas P.
; TITLE OF INVENTION: Nucleic Acids and Polypeptides Useful For Diagnosing and Treating
; FILE REFERENCE: 01948/108002
; CURRENT APPLICATION NUMBER: US/11/300,928
; CURRENT FILING DATE: 2005-12-15
; PRIOR APPLICATION NUMBER: US 60/636,275
; PRIOR FILING DATE: 2004-12-15
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-300-928-23

Query Match          0.2%; Score 18; DB 7; Length 561;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1030 TCCCAAGCAGTGGAGGA 1047
Db 342 TCCCAAGCAGTGGAGGA 359

RESULT 29
US-10-488-619-2900
; Sequence 2900, Application US/10488619
; Publication No. US20060099578A1
; GENERAL INFORMATION:
; APPLICANT: Greenlee, Winner and Sullivan, P. C.
; TITLE OF INVENTION: Human Mitochondrial DNA Polymorphisms, Haplogroups, Associations
; FILE REFERENCE: 98-01 WO
; CURRENT APPLICATION NUMBER: US/10/488,619
; CURRENT FILING DATE: 2004-03-01
; NUMBER OF SEQ ID NOS: 3040
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2900
; LENGTH: 575
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-488-619-2900

Query Match          0.2%; Score 18; DB 6; Length 575;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 ATGGGAGCTGTTTCATG 89
Db 430 ATGGGAGCTGTTTCATG 447

RESULT 30
US-11-218-305-24411

```

```

; Sequence 24411, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McIaird, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24411
; LENGTH: 840
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2)..(2)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5)..(5)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (14)..(14)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (777)..(777)
; OTHER INFORMATION: n is a, c, g, or t
US-11-218-305-24411

Query Match          0.2%; Score 18; DB 7; Length 840;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5687 AGTTTCTATGGAGATGC 5704
Db 136 AGTTTCTATGGAGATGC 153

Search completed: August 2, 2006, 09:18:12
Job time : 439 secs

```

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2006, 09:10:46 ; Search time 7477 Seconds

(without alignments)  
12460.191 Million cell updates/sec

Title: US-10-717-580-11

Perfect score: 7582

Sequence: 1 caacaatcggtataaacc.....tattaatactgcactgtg 7582

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 18892170 seqs, 6143817638 residues

Word size : 6

Total number of hits satisfying chosen parameters: 37734928

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 90 summaries

Database : Published Applications NA Main:\*

1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*

2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*

3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*

4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*

5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*

6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*

7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*

8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*

9: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*

10: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*

11: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10G\_PUBCOMB.seq:\*

12: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10H\_PUBCOMB.seq:\*

13: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10I\_PUBCOMB.seq:\*

14: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10J\_PUBCOMB.seq:\*

15: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11C\_PUBCOMB.seq:\*

16: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11D\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7582	100.0	7582	8	US-10-717-580-11
2	7582	100.0	7582	8	US-10-717-580-11
3	979	12.9	7466	10	US-10-450-763-4430
4	888	11.7	2372	8	US-10-717-580-11
5	888	11.7	2372	9	US-10-717-580-11
6	779	10.3	2599	13	US-11-028-539-1
7	779	10.3	8523	3	US-09-854-867-21
8	779	10.3	8523	10	US-10-786-970A-21
9	779	10.3	10499	13	US-11-028-539-3
10	779	10.3	56093	3	US-09-873-367C-81
11	779	10.3	56093	10	US-10-843-641A-81
12	750	9.9	6394	10	US-10-450-763-20014
13	731	9.6	1617	8	US-10-133-036-3
14	731	9.6	1948	8	US-10-632-793-24
15	731	9.6	1948	9	US-10-717-580-5
16	731	9.6	2055	13	US-11-028-539-22
17	731	9.6	2782	8	US-10-632-793-26

731	9.6	2782	8	US-10-133-036-1	Sequence 1, Appli
731	9.6	2782	9	US-10-717-580-7	Sequence 7, Appli
728	9.6	2930	3	US-09-902-535-1	Sequence 1, Appli
724	9.5	2938	8	US-10-632-793-21	Sequence 21, Appli
724	9.5	2938	8	US-10-717-580-2	Sequence 2, Appli
680	9.0	2946	6	US-10-114-893-134	Sequence 134, App
680	9.0	2946	6	US-10-016-249-3	Sequence 3, Appli
651	8.6	3633	10	US-10-821-234-543	Sequence 543, App
399	5.3	2006	8	US-10-632-793-23	Sequence 23, Appli
399	5.3	2006	9	US-10-717-580-4	Sequence 4, Appli
380	5.0	3372	8	US-10-450-763-20013	Sequence 20013, A
380	5.0	3372	9	US-10-632-793-28	Sequence 28, Appli
380	5.0	3372	9	US-10-717-580-9	Sequence 9, Appli
380	5.0	3372	10	US-10-450-763-7531	Sequence 7531, Ap
380	5.0	3372	10	US-10-450-763-3413	Sequence 3413, Ap
380	5.0	8279	10	US-10-450-763-12278	Sequence 12278, A
362	4.8	4349	10	US-10-450-763-12279	Sequence 12279, A
267	3.5	1326	13	US-11-028-539-2	Sequence 2, Appli
249	3.3	1136	8	US-10-632-793-25	Sequence 25, Appli
249	3.3	1136	9	US-10-717-580-6	Sequence 6, Appli
238	3.1	486	9	US-10-717-580-49	Sequence 49, Appli
213	2.8	1321	8	US-10-632-793-20	Sequence 20, Appli
213	2.8	1321	9	US-10-717-580-1	Sequence 1, Appli
213	2.8	1715	10	US-10-450-763-9505	Sequence 9505, Ap
211	2.8	600	9	US-10-717-580-47	Sequence 47, Appli
211	2.8	711	13	US-11-028-539-20	Sequence 20, Appli
211	2.8	873	9	US-10-717-580-43	Sequence 43, Appli
44	2.5	387	10	US-10-450-763-20006	Sequence 20006, A
46	2.5	542	6	US-10-040-916-48	Sequence 48, Appli
193	2.4	1080	13	US-11-028-539-28	Sequence 28, Appli
158	2.1	666	8	US-10-632-793-27	Sequence 27, Appli
158	2.1	666	9	US-10-717-580-8	Sequence 8, Appli
157	2.1	910	10	US-10-450-763-7529	Sequence 7529, Ap
140	1.8	530	9	US-10-717-580-48	Sequence 48, Appli
138	1.8	783	9	US-10-717-580-15	Sequence 15, Appli
116	1.5	629	12	US-10-301-480-531649	Sequence 531649, Sequence 1205058,
116	1.5	629	12	US-10-301-480-1205058	Sequence 1205058,
104	1.4	893	10	US-10-450-763-1768	Sequence 1768, Ap
89	1.2	425	9	US-10-717-580-45	Sequence 45, Appli
89	1.2	427	9	US-10-717-580-46	Sequence 46, Appli
89	1.2	711	13	US-11-028-539-21	Sequence 21, Appli
85	1.2	815	9	US-10-717-580-44	Sequence 44, Appli
85	1.2	2074	8	US-10-416-642-4	Sequence 4, Appli
61	1.1	630	4	US-09-925-065A-739053	Sequence 739053,
62	1.1	630	5	US-09-925-065A-739053	Sequence 739053,
63	1.1	631	4	US-09-925-065A-755821	Sequence 755821,
82	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
82	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
64	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
65	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
65	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
66	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
67	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
68	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
68	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
69	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
70	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
71	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
72	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
73	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
74	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
75	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
76	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
77	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
78	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
79	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
80	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
81	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
82	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
83	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
84	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
85	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
86	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
87	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
88	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
89	1.1	631	5	US-09-925-065A-755821	Sequence 755821,
90	1.1	631	5	US-09-925-065A-755821	Sequence 755821,

ALIGNMENTS

```

RESULT 1
US-10-632-793-30
; Sequence 30, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: MALLET, Francois
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
; FILE REFERENCE: 110048
; CURRENT APPLICATION NUMBER: US/10/632,793
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/FR00/00144
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 7582
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (198)..(198)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (307)..(307)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (355)..(355)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1309)..(1309)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1331)..(1331)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2213)..(2213)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2398)..(2398)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3787)..(3787)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4115)..(4115)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4261)..(4261)
; OTHER INFORMATION: n = a or g or c or t/u
US-10-632-793-30

```

Query Match 100.0%; Score 7582; DB 8; Length 7582;

		Best Local Similarity 100.0%; Pred. No. 0;				Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	DB								
1	1	CAACAATCGGATATAAACCAGGCAATTCGAGCTGGGCAACAGAGCCGCCCTTTGGGTCC		60					
61	61	CTTCCCTTGTATGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCACCTGCA		120					
121	121	CTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC		180					
181	181	TGTTTGGCCACACCGCANACCTGCGGTGACTCCCATCCCTCTGGATCCTGCAGGGTGC		240					
301	301	CATTTGNCCTGCACGGCTTAAGTGCCTGGTGTGTTCTAATTTGAGCTGAACACTANTCACT		360					
421	421	TGGCCCAAGATTCCATTCCTTTGGAATCCGTGAGGSCAAGAACTCCAGGTAGAGAATAC		480					
541	541	TGGGAGCTCTGTGAGCAAGGACCCCGGTGACATTTTGGCRACCAAGGACATCC		600					
661	661	TGGGAGCTCTGTGAGCAAGGACCCCGGTGACATTTTGGCRACCAAGGACATCC		660					
781	781	TAGAAAAAGGCAAAATGGAGTGAAGTGCCTAAGTACAAACTTTTCTTTTCAATTAAGACAA		840					
901	901	CTATCCAGCATCCCGGACTCTTCCCGGACTCTTCCCGGACTCTTCCCGGACTCTTCCCGG		960					
1021	1021	TATGACCCCTCCCAAGCAGTGGAGAGAGAAATTCGGCCCGCAGGAGTGCATGTCYT		1080					



Db 1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAATTCCGCCAGCCAGAGTGCATGTGCT 1080  
Qy 1081 TTTTTCCTCCAGACTTAAAGCAATAAATAAAGCAAGACTTAGGTAAATCTCAGATAAYCCT 1140  
Db 1081 TTTTTCCTCCAGACTTAAAGCAATAAATAAAGCAAGACTTAGGTAAATCTCAGATAAYCCT 1140  
Qy 1141 GATGGCTATATGTGTGTTTACAAGGGTTAGGACAATTTCTTGATCTGACATGGAGAGAT 1200  
Db 1141 GATGGCTATATGTGTGTTTACAAGGGTTAGGACAATTTCTTGATCTGACATGGAGAGAT 1200  
Qy 1201 ATATATGTCTGCTTAATCAGACACTAAACCCCAAAATGAGAGAAGTGCCACCAATACTGC 1260  
Db 1201 ATATATGTCTGCTTAATCAGACACTAAACCCCAAAATGAGAGAAGTGCCACCAATACTGC 1260  
Qy 1261 AGCTCAGAGTGTGGCGATCTCTGTTATCTCAGTCAGTCAATGGATGAGATGACACACA 1320  
Db 1261 AGCTCAGAGTGTGGCGATCTCTGTTATCTCAGTCAGTCAATGGATGAGATGACACACA 1320  
Qy 1321 GAAGGAAGANAATGATTTCCCAAGCCAGCAGCAGTTCGCCAGTCTASACCTCATTTG 1380  
Db 1321 GAAGGAAGANAATGATTTCCCAAGCCAGCAGCAGTTCGCCAGTCTASACCTCATTTG 1380  
Qy 1381 GGGACACAGAAATCAGTAAATCGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGTGC 1440  
Db 1381 GGGACACAGAAATCAGTAAATCGGAGATTGGTGCTGCAGACATTTGCTAACTTGTGTGC 1440  
Qy 1441 TASAAGGACTAAGGAATACTASAGAANAATCTAYGAATTAATCAATGATGTCACACATA 1500  
Db 1441 TASAAGGACTAAGGAATACTASAGAANAATCTAYGAATTAATCAATGATGTCACACATA 1500  
Qy 1501 ACACAGGGGAAGGAAGAAATCCTACTGCTTCTGAGGAGACTAAGGGAGGCAATTGAG 1560  
Db 1501 ACACAGGGGAAGGAAGAAATCCTACTGCTTCTGAGGAGACTAAGGGAGGCAATTGAG 1560  
Qy 1561 GAAGCGTGCTCTCTGCTCAGCTGACTTCTGAAGGCCCAACTAATCTTAAAGCGTAAGTT 1620  
Db 1561 GAAGCGTGCTCTCTGCTCAGCTGACTTCTGAAGGCCCAACTAATCTTAAAGCGTAAGTT 1620  
Qy 1621 TATCACTCAGTGCAGACATTAAGAAAACTTCAAAAGTCTGCCGTAGGCCCGGAG 1680  
Db 1621 TATCACTCAGTGCAGACATTAAGAAAACTTCAAAAGTCTGCCGTAGGCCCGGAG 1680  
Qy 1681 CAAAACCTTAAAGCCCTATTTGAATTTGGCACTTGGCACTTGGCACTTGGCACTTGGCA 1740  
Db 1681 CAAAACCTTAAAGCCCTATTTGAATTTGGCACTTGGCACTTGGCACTTGGCACTTGGCA 1740  
Qy 1741 GAGCAGCGGAAACAGGACAAACGGGATTAAGAAAAAGGCAAGCCGCTTTAGTCATGACCT 1800  
Db 1741 GAGCAGCGGAAACAGGACAAACGGGATTAAGAAAAAGGCAAGCCGCTTTAGTCATGACCT 1800  
Qy 1801 CAGGCAAGTGGACTTTGGAGGCTCTGAAAAGGGNAAGCTGGGCAAAATGGAATGCTTAA 1860  
Db 1801 CAGGCAAGTGGACTTTGGAGGCTCTGAAAAGGGNAAGCTGGGCAAAATGGAATGCTTAA 1860  
Qy 1861 TAGGGCTTGCCTCCAGTGGGTCTACAAGGACATTTTAAAGATGTGTCAGATAGAAG 1920  
Db 1861 TAGGGCTTGCCTCCAGTGGGTCTACAAGGACATTTTAAAGATGTGTCAGATAGAAG 1920  
Qy 1921 TAAGCGCCCTCTGCTGATGCCCTTATTTCAAGGGAATCACTGGAAGCCCACTGCC 1980  
Db 1921 TAAGCGCCCTCTGCTGATGCCCTTATTTCAAGGGAATCACTGGAAGCCCACTGCC 1980  
Qy 1981 CAGGGGCAAAAGTCTTTTGAAGTCAAGGACCACTAACAGATGATCCAGCAGGAGCTG 2040  
Db 1981 CAGGGGCAAAAGTCTTTTGAAGTCAAGGACCACTAACAGATGATCCAGCAGGAGCTG 2040  
Qy 2041 AGGGTCCCTGGGCAAGGCCATCCATGCCATCACCCTCAGAGGCCCTGGGTATGCTT 2100  
Db 2041 AGGGTCCCTGGGCAAGGCCATCCATGCCATCACCCTCAGAGGCCCTGGGTATGCTT 2100  
Qy 2101 GACCATTTAGGGCCAGAGGTTGTCTCTGACACTGTGGCTTCTTAGTCTTACT 2160  
Db 2101 GACCATTTAGGGCCAGAGGTTGTCTCTGACACTGTGGCTTCTTAGTCTTACT 2160

Qy 2161 CTTCTGTCCGGACAACTGTCTCTCCAGATCTGTCACTATTCTGAGGGGTCNTAAGACG 2220  
Db 2161 CTTCTGTCCGGACAACTGTCTCTCCAGATCTGTCACTATTCTGAGGGGTCNTAAGACG 2220  
Qy 2221 GGCAGTCACTAGANACTTTTCCAGGCACTAAGTTATGAACCTGGGAGCTTTATCTTT 2280  
Db 2221 GGCAGTCACTAGANACTTTTCCAGGCACTAAGTTATGAACCTGGGAGCTTTATCTTT 2280  
Qy 2281 TCACATGCTTTTCTAATATATGCTTGAAGCCCACTACCTTTGTTAGGGAGACATTTCTA 2340  
Db 2281 TCACATGCTTTTCTAATATATGCTTGAAGCCCACTACCTTTGTTAGGGAGACATTTCTA 2340  
Qy 2341 GCMAAGCAGGGGCCATTTATACACTGAACAATAGGAGAGGAACAACCCGTTTGTGTNCC 2400  
Db 2341 GCMAAGCAGGGGCCATTTATACACTGAACAATAGGAGAGGAACAACCCGTTTGTGTNCC 2400  
Qy 2401 CTTGCTTTGAGGAGGAAGTAATTCCTGAAGTCTGGGCAACAGAGGACATATGAGCGAGC 2460  
Db 2401 CTTGCTTTGAGGAGGAAGTAATTCCTGAAGTCTGGGCAACAGAGGACATATGAGCGAGC 2460  
Qy 2461 CAAAGAAATGCCGCTCCTGTTCAAGTTAAACTAAAGGATTCACCTTCTTCCCTACAAA 2520  
Db 2461 CAAAGAAATGCCGCTCCTGTTCAAGTTAAACTAAAGGATTCACCTTCTTCCCTACAAA 2520  
Qy 2521 GGGAGTACCCCTCAGACCCAAAGGCCAAAGGATTCAAAAGATTTGTTAAGGACTTAA 2580  
Db 2521 GGGAGTACCCCTCAGACCCAAAGGCCAAAGGATTCAAAAGATTTGTTAAGGACTTAA 2580  
Qy 2581 AAGCCCAAGGCTTACTTAAACCATGCACTCCCTGCGAGTAATTCCTGAGTGGATTGAG 2640  
Db 2581 AAGCCCAAGGCTTACTTAAACCATGCACTCCCTGCGAGTAATTCCTGAGTGGATTGAG 2640  
Qy 2641 GAGGACACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAGAGATCTCAGGATTAATCAATG 2700  
Db 2641 GAGGACACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAGAGATCTCAGGATTAATCAATG 2700  
Qy 2701 AGGCGTGTGCTTTTATACCCAGCTGTACTAGCCCTTATCTGTGMYTTTCCCAATATC 2760  
Db 2701 AGGCGTGTGCTTTTATACCCAGCTGTACTAGCCCTTATCTGTGMYTTTCCCAATATC 2760  
Qy 2761 CAGAGGACAGAGTGGTTTACSTCTGGAACCTTMAAGATGCTTCTTCTGATCCCTG 2820  
Db 2761 CAGAGGACAGAGTGGTTTACSTCTGGAACCTTMAAGATGCTTCTTCTGATCCCTG 2820  
Qy 2821 TACATCTGACTCTCAATTTCTTTGTTGCTTTGAAAGATACCTTCAAAACCCAGATCTCAAC 2880  
Db 2821 TACATCTGACTCTCAATTTCTTTGTTGCTTTGAAAGATACCTTCAAAACCCAGATCTCAAC 2880  
Qy 2881 TCACCTGGACTTTTAAACCCAGGTTTCAAGGATAGYCCCATCTATTTGGCCAGGCAT 2940  
Db 2881 TCACCTGGACTTTTAAACCCAGGTTTCAAGGATAGYCCCATCTATTTGGCCAGGCAT 2940  
Qy 2941 TAGCCCAAGACTTGAGYCAATYMTATACCTGGAACCTTCTGCTCTGTAAGTGGATG 3000  
Db 2941 TAGCCCAAGACTTGAGYCAATYMTATACCTGGAACCTTCTGCTCTGTAAGTGGATG 3000  
Qy 3001 ATTTACTTTTTCGCTTTCAGAAACCTTGTGTCATCAAGCCCAAGGCTCTTTMA 3060  
Db 3001 ATTTACTTTTTCGCTTTCAGAAACCTTGTGTCATCAAGCCCAAGGCTCTTTMA 3060  
Qy 3061 ATTTCTCTGCTGCTGCTTACAGGTTTCCAAACSAARAGCTCARTCTCTGCTCACAGC 3120  
Db 3061 ATTTCTCTGCTGCTGCTTACAGGTTTCCAAACSAARAGCTCARTCTCTGCTCACAGC 3120  
Qy 3121 AGGTAAATATCTAGGCTTAARATATCAAAAGGACCAAGGCCCTCAGTGGAGAAATYA 3180  
Db 3121 AGGTAAATATCTAGGCTTAARATATCAAAAGGACCAAGGCCCTCAGTGGAGAAATYA 3180  
Qy 3181 TCCAGCTTACTGCTTATCTCTATCYCAAAACCTTAAGCAACTAAGGRRTTCTTGT 3240  
Db 3181 TCCAGCTTACTGCTTATCTCTATCYCAAAACCTTAAGCAACTAAGGRRTTCTTGT 3240

QY 3241 GCRTAAYAGGYTTCTGCCGAAWATGGATTCCCGAGTGTGGCRAAATAGCCAGGYCATTA 3300  
Db |||||  
QY 3241 GCRTAAYAGGYTTCTGCCGAAWATGGATTCCCGAGTGTGGCRAAATAGCCAGGYCATTA 3300  
Db |||||  
QY 3301 WATACASTAATTAAGGAACTCAGAAAGCCCAATACCCATTARTTAAGATGGAYAMCTGAA 3360  
Db |||||  
QY 3301 WATACASTAATTAAGGAACTCAGAAAGCCCAATACCCATTARTTAAGATGGAYAMCTGAA 3360  
Db |||||  
QY 3361 GYMAAGTGGCTTTCAGGCCCTTAAAGAGGCCCTTAAACCAAGYCCAGGTGTTAAGYT 3420  
Db |||||  
QY 3361 GYMAAGTGGCTTTCAGGCCCTTAAAGAGGCCCTTAAACCAAGYCCAGGTGTTAAGYT 3420  
Db |||||  
QY 3421 TGCCAACRGGGCAAGACTTTTSTTYATAYRTACAGAAAAAACAAGRAYAGCTCTRGA 3480  
Db |||||  
QY 3481 GTCCCTTACACAGRTCCRAGGGAYGAGCTTGCAACCCYRTGGCRYACCTGAAGGAAAYT 3540  
Db |||||  
QY 3541 GATGTAGTGGCAAGGGTGTGCYTCATTTGTTAYGGGTAGTGGTGGCAGTAGCAGTYKTA 3600  
Db |||||  
QY 3541 GATGTAGTGGCAAGGGTGTGCYTCATTTGTTAYGGGTAGTGGTGGCAGTAGCAGTYKTA 3600  
Db |||||  
QY 3601 GTATCTGAAGCAGTTTAAATAATACAGGGRAGAGATCTTACTGTGTGCACATCTCATGAK 3660  
Db |||||  
QY 3601 GTATCTGAAGCAGTTTAAATAATACAGGGRAGAGATCTTACTGTGTGCACATCTCATGAK 3660  
Db |||||  
QY 3661 GTGAAYRGCACTCACTGCTCTAAAGGAGACTTGTGGCTGTGCAGAAACCTTAA 3720  
Db |||||  
QY 3721 TRTCAAGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACCTTCTTAAC 3780  
Db |||||  
QY 3721 TRTCAAGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACCTTCTTAAC 3780  
Db |||||  
QY 3781 CCAGYCNCAATTTCTCCAGACAAATGAAGAAAGATARAAYATACTGTCAACAARTTAAT 3840  
Db |||||  
QY 3781 CCAGYCNCAATTTCTCCAGACAAATGAAGAAAGATARAAYATACTGTCAACAARTTAAT 3840  
Db |||||  
QY 3841 TCTCAAAACCTATGCCACTCGAGGGGACCTTGTAGARGTTTCCTTGAATCCYAGCCTT 3900  
Db |||||  
QY 3901 CAACTGTGATGATGAGAGTTCCTTTGTAGAAAAAGACTTCGAAAGYGGGTATGC 3960  
Db |||||  
QY 3901 CAACTGTGATGATGAGAGTTCCTTTGTAGAAAAAGACTTCGAAAGYGGGTATGC 3960  
Db |||||  
QY 3961 AGTGGTCACTGATGATAATGGAATAYTTGAAAGTAATCCCTCACTCCAGGAACCTAGTCTYA 4020  
Db |||||  
QY 3961 AGTGGTCACTGATGATAATGGAATAYTTGAAAGTAATCCCTCACTCCAGGAACCTAGTCTYA 4020  
Db |||||  
QY 4021 GCTRCAGAACCTAATAGCCCTCACTKGGGCACTAGAAATAGGAGAAAGAAAGGGYAAA 4080  
Db |||||  
QY 4021 GCTRCAGAACCTAATAGCCCTCACTKGGGCACTAGAAATAGGAGAAAGAAAGGGYAAA 4080  
Db |||||  
QY 4081 TATATATACAGCTCTTRARTATGCTYACCTAGTGTGTCATGTCCTGMRGCAATATGSAR 4140  
Db |||||  
QY 4081 TATATATACAGCTCTTRARTATGCTYACCTAGTGTGTCATGTCCTGMRGCAATATGSAR 4140  
Db |||||  
QY 4141 AGAAGGGGAATTCCTAACTTCYAGRGAACTATCAMACATCAGGAAGCCATTAGGAR 4200  
Db |||||  
QY 4141 AGAAGGGGAATTCCTAACTTCYAGRGAACTATCAMACATCAGGAAGCCATTAGGAR 4200  
Db |||||  
QY 4201 ATTATATTATGCGWGTACAGAAACCTTARAGAGGTGGMAGTCTTACACTGCGGGGTCACTA 4260  
Db |||||  
QY 4201 ATTATATTATGCGWGTACAGAAACCTTARAGAGGTGGMAGTCTTACACTGCGGGGTCACTA 4260  
Db |||||  
QY 4261 NAAAGGAAAGRAAGGAAATASAGRGAAAYTGCCAAAGCAKATATTGAAGCMAAAAGAGC 4320  
Db |||||  
QY 4261 NAAAGGAAAGRAAGGAAATASAGRGAAAYTGCCAAAGCAKATATTGAAGCMAAAAGAGC 4320  
Db |||||  
QY 4321 TGCAAGGCGAGGACCTCCATTAGAAATGCTTATTAACCTTCCCTTAGTATAGGGTAATCC 4380

Db |||||  
QY 4321 TGCAAGGCGAGGACCTCCATTAGAAATGCTTATTAACCTTCCCTTAGTATAGGGTAATCC 4380  
Db |||||  
QY 4381 CTTCCGGGAAACCAAGCCCTCAGTAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCACAGG 4440  
Db |||||  
QY 4381 CTTCCGGGAAACCAAGCCCTCAGTAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCACAGG 4440  
Db |||||  
QY 4441 CAGTTTTCTCCCTCGGGACGGTTAGCCACTGAAAGAGGGGAAATATCTTTTGCCTGCAC 4500  
Db |||||  
QY 4441 CAGTTTTCTCCCTCGGGACGGTTAGCCACTGAAAGAGGGGAAATATCTTTTGCCTGCAC 4500  
Db |||||  
QY 4501 TATCAATGGAATTAATCTTAAAAACCTTCAATCAAAACCTTTCATAGGATCGATAGCAC 4560  
Db |||||  
QY 4501 TATCAATGGAATTAATCTTAAAAACCTTCAATCAAAACCTTTCATAGGATCGATAGCAC 4560  
Db |||||  
QY 4561 CCATCARATGGCCCAATCATTTATTTACTTGGACAGGCCCTTTTCAAACTATCAAGCARAT 4620  
Db |||||  
QY 4561 CCATCARATGGCCCAATCATTTATTTACTTGGACAGGCCCTTTTCAAACTATCAAGCARAT 4620  
Db |||||  
QY 4621 AKTCAGGCCCTGTGAATGTGCCARARAAATAATCCCTGCTCTATCCCAAGCTCTTC 4680  
Db |||||  
QY 4621 AKTCAGGCCCTGTGAATGTGCCARARARAAATAATCCCTGCTCTATCCCAAGCTCTTC 4680  
Db |||||  
QY 4681 AGGARAACAAARAAACAGGCCATTACCTTGRARAARACTGGCAACTGATTTTACCCACAAG 4740  
Db |||||  
QY 4681 AGGARAACAAARAAACAGGCCATTACCTTGRARAARACTGGCAACTGATTTTACCCACAAG 4740  
Db |||||  
QY 4741 CCCAAACCTCAGGATTTAGTATCTACTAGTCTGGGTARATATCTTTCACGGTGGGCA 4800  
Db |||||  
QY 4741 CCCAAACCTCAGGATTTAGTATCTACTAGTCTGGGTARATATCTTTCACGGTGGGCA 4800  
Db |||||  
QY 4801 RAGGCTTTCCCTGTAGGACAGAAAGGCCCAAGAGTAATAAAGGCACTAGTTTCACTGAA 4860  
Db |||||  
QY 4801 RAGGCTTTCCCTGTAGGACAGAAAGGCCCAAGAGTAATAAAGGCACTAGTTTCACTGAA 4860  
Db |||||  
QY 4861 ATAAATCCAGATTTCCGACTTTCCCGAGGCTTACAGAGTGACAAATAGCCCTTCCAG 4920  
Db |||||  
QY 4861 ATAAATCCAGATTTCCGACTTTCCCGAGGCTTACAGAGTGACAAATAGCCCTTCCAG 4920  
Db |||||  
QY 4921 GCCACAGTAACCCAGGAGTATCCAGCGTTAGGTATACGATATCACTTACACTGCGCC 4980  
Db |||||  
QY 4921 GCCACAGTAACCCAGGAGTATCCAGCGTTAGGTATACGATATCACTTACACTGCGCC 4980  
Db |||||  
QY 4981 TGAAGGCCACAGTCTCCTCAGGGAAGGTGAGAAATGAATGAAAYACTCAAGGACATCTA 5040  
Db |||||  
QY 4981 TGAAGGCCACAGTCTCCTCAGGGAAGGTGAGAAATGAATGAAAYACTCAAGGACATCTA 5040  
Db |||||  
QY 5041 AAAAGGCAAAACCCAGGAAACCCCACTCAGTGGCTGTCTGTGGCTTATAGCCCTTAAAA 5100  
Db |||||  
QY 5041 AAAAGGCAAAACCCAGGAAACCCCACTCAGTGGCTGTCTGTGGCTTATAGCCCTTAAAA 5100  
Db |||||  
QY 5101 AGAATCTGCAACTTTTCCCAAAAAGCAGGACTTAGCCCATAGAAATGCTGTATGGAAG 5160  
Db |||||  
QY 5101 AGAATCTGCAACTTTTCCCAAAAAGCAGGACTTAGCCCATAGAAATGCTGTATGGAAG 5160  
Db |||||  
QY 5161 CCCTTCATAACCAATGACTTGTGCTGACCCCAAGACCACTTAGTGTGAGACATCA 5220  
Db |||||  
QY 5161 CCCTTCATAACCAATGACTTGTGCTGACCCCAAGACCACTTAGTGTGAGACATCA 5220  
Db |||||  
QY 5221 CCTCCTTAGCCAAATATCAACAAGTCTTAAAAANTTACAGGAACCTTATCCCTGGAAG 5280  
Db |||||  
QY 5281 AGGGAAGAAAGACTATTCCACCWGTGACATGTTAGTCAAGTCCCTTCTCTAAAT 5340  
Db |||||  
QY 5341 CCCCATCCCTAGATACATCTCTGGGAGGACCTTACCCAGTCAATTTATTTTACCCCACTG 5400  
Db |||||  
QY 5341 CCCCATCCCTAGATACATCTCTGGGAGGACCTTACCCAGTCAATTTATTTTACCCCACTG 5400  
Db |||||  
QY 5401 CGGTTAAAGTGGCTGGAGTCTTGGATACATCACACTTGAGTCAAAATCCTGCGATAC 5460

Db 5401 CGGTTAAAGTGGCTGGAGTCTTGATACATCACTTGAGTCAAAATCCTGGATAC 5460  
Qy 5461 TGCCAAAGGAACCTGAAATCCAGAGACAAAGCTAGCTATTCCTGTGAACCTCTAGAGG 5520  
Db 5461 TGCCAAAGGAACCTGAAATCCAGAGACAAAGCTAGCTATTCCTGTGAACCTCTAGAGG 5520  
Qy 5521 ATTTCGGCCTGCTCTTCAAAACAAACACGAGGAGAAAGTAACTAAATCATAAATCCCCC 5580  
Db 5521 ATTTCGGCCTGCTCTTCAAAACAAACACGAGGAGAAAGTAACTAAATCATAAATCCCCC 5580  
Qy 5581 ATGSGCCTCCCTTATCATATTTTTCTCTKTAAGTCTTTTACCCCTSTTTCACTCTCACT 5640  
Db 5581 ATGSGCCTCCCTTATCATATTTTTCTCTKTAAGTCTTTTACCCCTSTTTCACTCTCACT 5640  
Qy 5641 GCACCCCTCATGCGCTGTATGACCAAGTAGCTCCCTTACCMAGAGTTTCTATGAGGA 5700  
Db 5641 GCACCCCTCATGCGCTGTATGACCAAGTAGCTCCCTTACCMAGAGTTTCTATGAGGA 5700  
Qy 5701 ATGCAGCGTCCCGGAAATATTGATGCCCAATCGTATAGGAGTCTTTSTAAGGGAACCCC 5760  
Db 5701 ATGCAGCGTCCCGGAAATATTGATGCCCAATCGTATAGGAGTCTTTSTAAGGGAACCCC 5760  
Qy 5761 ACCTTCACTGCCACACCCATATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCAATG 5820  
Db 5761 ACCTTCACTGCCACACCCATATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCAATG 5820  
Qy 5821 CATGCCAAATCACTATTTATGGAAGGAAATATTAATCCTAGTTGTCTGGAGGACTT 5880  
Db 5821 CATGCCAAATCACTATTTATGGAAGGAAATATTAATCCTAGTTGTCTGGAGGACTT 5880  
Qy 5881 GGAGTCACTGTCTGTGGACTTACTTCAACCAACTGCTATGCTGATGGGGTGGAGTT 5940  
Db 5881 GGAGTCACTGTCTGTGGACTTACTTCAACCAACTGCTATGCTGATGGGGTGGAGTT 5940  
Qy 5941 CAAGATCAGGCAAGAGAAAACATGTAAAGAGTAATCTCCCAACTCAACGSGGTACAT 6000  
Db 5941 CAAGATCAGGCAAGAGAAAACATGTAAAGAGTAATCTCCCAACTCAACGSGGTACAT 6000  
Qy 6001 GGACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTAAC 6060  
Db 6001 GGACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTAAC 6060  
Qy 6061 CATACTCGCTGGTAAAGCTATTTAATACACCCCTCACTGGGCTCATGAGGCTCGGCC 6120  
Db 6061 CATACTCGCTGGTAAAGCTATTTAATACACCCCTCACTGGGCTCATGAGGCTCGGCC 6120  
Qy 6121 CAAACCCCTACTAATCTGTGGATGCTCCCTCGAACTTTCARGCCATATGTTCAATC 6180  
Db 6121 CAAACCCCTACTAATCTGTGGATGCTCCCTCGAACTTTCARGCCATATGTTCAATC 6180  
Qy 6181 CTTGTACTGAACAAATGAACAACTTCAGCACAGAAATAAACACCACTTCCTGTTTTAGTA 6240  
Db 6181 CTTGTACTGAACAAATGAACAACTTCAGCACAGAAATAAACACCACTTCCTGTTTTAGTA 6240  
Qy 6241 GGACCTCTGTTTCCAAATSTGGAATTAACCACTCACTCAAACTCACTGTGTAAATTT 6300  
Db 6241 GGACCTCTGTTTCCAAATSTGGAATTAACCACTCACTCAAACTCACTGTGTAAATTT 6300  
Qy 6301 AGCAATACTACATACAAACCACTCCCAATGCACTGAGTGGGTAACTCTCCACACAA 6360  
Db 6301 AGCAATACTACATACAAACCACTCCCAATGCACTGAGTGGGTAACTCTCCACACAA 6360  
Qy 6361 ATAGTCTGCTTACCTCAGGAATATTTTTGTCTGTGGTACCTCAGGCTATCGTTGTTTG 6420  
Db 6361 ATAGTCTGCTTACCTCAGGAATATTTTTGTCTGTGGTACCTCAGGCTATCGTTGTTTG 6420  
Qy 6421 AATGGCTCTTCAAGATCTATGTGCTTCTCTCAATCTTCTAGTGGCCCAATGRCATCTAC 6480  
Db 6421 AATGGCTCTTCAAGATCTATGTGCTTCTCTCAATCTTCTAGTGGCCCAATGRCATCTAC 6480  
Qy 6481 ACTGACAAAGATTTATACAGTTATGTCATATCTAAGCCCGGCAACAAAGAGTACCCATT 6540  
Db 6481 ACTGACAAAGATTTATACAGTTATGTCATATCTAAGCCCGGCAACAAAGAGTACCCATT 6540

Qy 6541 CTTCCCTTTTGTATAGGAGCAGAGTGTCTAGGTGCACCTAGGTACTGCGATTTGGCGGTATC 6600  
Db 6541 CTTCCCTTTTGTATAGGAGCAGAGTGTCTAGGTGCACCTAGGTACTGCGATTTGGCGGTATC 6600  
Qy 6601 ACACCTCTACTCAGTTTCTACTCAAACTATCTCAAGAACTAAATGGGACATGGAACGG 6660  
Db 6601 ACACCTCTACTCAGTTTCTACTCAAACTATCTCAAGAACTAAATGGGACATGGAACGG 6660  
Qy 6661 GTGCCGACTCCCTGGTCACTTGCAGAGTCAAACTCCCTAGCAGAGTAGTCTCTT 6720  
Db 6661 GTGCCGACTCCCTGGTCACTTGCAGAGTCAAACTCCCTAGCAGAGTAGTCTCTT 6720  
Qy 6721 CRAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAGAGGGGAACTGTATTTTTA 6780  
Db 6721 CRAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAGAGGGGAACTGTATTTTTA 6780  
Qy 6781 GGGGAGAAATGCTGTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 6840  
Db 6781 GGGGAGAAATGCTGTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 6840  
Qy 6841 ATTCSAGATCGAATAAACAAGTAKAGCAGAGAGCTTCAAACTAGGACCTCTGGGGCTC 6900  
Db 6841 ATTCSAGATCGAATAAACAAGTAKAGCAGAGAGCTTCAAACTAGGACCTCTGGGGCTC 6900  
Qy 6901 CTGAGCCRAATGAGTCCCTGGATTTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATATTG 6960  
Db 6901 CTGAGCCRAATGAGTCCCTGGATTTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATATTG 6960  
Qy 6961 CTACTCTCTTTGGACCCCTGTATCTTTTACCTCTTTTAACTTTGTCTCTCCAGAACTC 7020  
Db 6961 CTACTCTCTTTGGACCCCTGTATCTTTTACCTCTTTTAACTTTGTCTCTCCAGAACTC 7020  
Qy 7021 GAAGCTGTAAACTCAAAATGGAGCCCAAGTCCAGTCCAAAGTAACTAAGTCTACCCGAGA 7080  
Db 7021 GAAGCTGTAAACTCAAAATGGAGCCCAAGTCCAGTCCAAAGTAACTAAGTCTACCCGAGA 7080  
Qy 7081 CCCCTGGACCGGCTGTAGCCCAAGTCTGATGTTTAAATGACATCAAGGACACCCCTCTCT 7140  
Db 7081 CCCCTGGACCGGCTGTAGCCCAAGTCTGATGTTTAAATGACATCAAGGACACCCCTCTCT 7140  
Qy 7141 GAGGAATCTCAGCTGCACAACTCTTAACGCCCCCAATTCAGAGAGAGCAGTTAGAGC 7200  
Db 7141 GAGGAATCTCAGCTGCACAACTCTTAACGCCCCCAATTCAGAGAGAGCAGTTAGAGC 7200  
Qy 7201 GGTSGTGGCCCACTCCCAAGCAGACTTAGGTTTCTGTTGAGTGGGAGCTGAGA 7260  
Db 7201 GGTSGTGGCCCACTCCCAAGCAGACTTAGGTTTCTGTTGAGTGGGAGCTGAGA 7260  
Qy 7261 GACAGGACTAGCTGGATTTCTTAGGCTGATTAAGAACTCCYTAAGCCTTAGTGGAAGGTG 7320  
Db 7261 GACAGGACTAGCTGGATTTCTTAGGCTGATTAAGAACTCCYTAAGCCTTAGTGGAAGGTG 7320  
Qy 7321 ACCAATCAACCTTTTAAACACGCGGCTTGCACTTAGYTCAACCTGACCAATCAGAGAG 7380  
Db 7321 ACCAATCAACCTTTTAAACACGCGGCTTGCACTTAGYTCAACCTGACCAATCAGAGAG 7380  
Qy 7381 CTCACCTAAATGCTAATTTAGGCAAGACAGAGGTTAAAGAAATAGCAATCATYTTATTC 7440  
Db 7381 CTCACCTAAATGCTAATTTAGGCAAGACAGAGGTTAAAGAAATAGCAATCATYTTATTC 7440  
Qy 7441 MTGAGAGCAGCAGCAGGAGGCAATGATCGGATATAAACCAGYTTTCGAGCCCGCAA 7500  
Db 7441 MTGAGAGCAGCAGCAGGAGGCAATGATCGGATATAAACCAGYTTTCGAGCCCGCAA 7500  
Qy 7501 CGGCAACCCCTTTGGGTCCCTCTTTGTATGGAGCTCTGTTTTCACTATTTTCAC 7560  
Db 7501 CGGCAACCCCTTTGGGTCCCTCTTTGTATGGAGCTCTGTTTTCACTATTTTCAC 7560  
Qy 7561 TCTATTAATCTTGCARCTGCR 7582  
Db 7561 TCTATTAATCTTGCARCTGCR 7582

RESULT 2  
US-10-717-580-11  
; Sequence 11, Application US/10717580  
; Publication No. US20040176314A1  
; GENERAL INFORMATION:  
; APPLICANT: BESEME, Frederic  
; APPLICANT: BLOND, Jean-Luc  
; APPLICANT: BOUTON, Olivier  
; APPLICANT: MANDRAND, Bernard  
; APPLICANT: MALLET, Francois  
; APPLICANT: PERRON, Herve  
; TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISEASES  
; TITLE OF INVENTION: WITH PREGNANCY DISORDERS  
; FILE REFERENCE: 105045  
; CURRENT APPLICATION NUMBER: US/10/717,580  
; CURRENT FILING DATE: 2003-11-21  
; PRIOR APPLICATION NUMBER: US/09/446,024A  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: PCT/FR98/01442  
; PRIOR FILING DATE: 1998-07-06  
; PRIOR APPLICATION NUMBER: FR 97/08815  
; PRIOR FILING DATE: 1997-07-07  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 11  
; LENGTH: 7582  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (198)..(198)  
; OTHER INFORMATION: n = any nucleotide  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (307)..(307)  
; OTHER INFORMATION: n = any nucleotide  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (355)..(355)  
; OTHER INFORMATION: n = any nucleotide  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1309)..(1309)  
; OTHER INFORMATION: n = any nucleotide  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1331)..(1331)  
; OTHER INFORMATION: n = any nucleotide  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (2213)..(2213)  
; OTHER INFORMATION: n = any nucleotide  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (2398)..(2398)  
; OTHER INFORMATION: n = any nucleotide  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (3787)..(3787)  
; OTHER INFORMATION: n = any nucleotide  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (4115)..(4115)  
; OTHER INFORMATION: n = any nucleotide  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (4261)..(4261)  
; OTHER INFORMATION: n = any nucleotide  
; US-10-717-580-11

Query Match 100.0%; Score 7582; DB 9; Length 7582;  
Best Local Similarity 100.0%; Pred. No. 0;

	Matches	7582;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	CAACAATCGGATATATAAACCCAGGCAATTCGAGCTGGGCAACAGCAGCCCTTTGGGTCC	60							
DB	1	CAACAATCGGATATATAAACCCAGGCAATTCGAGCTGGGCAACAGCAGCCCTTTGGGTCC	60							
QY	61	CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACCTCTATTAAATCTTGCACACTGCA	120							
DB	61	CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACCTCTATTAAATCTTGCACACTGCA	120							
QY	121	CTCTTCTGGTCCATGTTTCTTAAGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC	180							
DB	121	CTCTTCTGGTCCATGTTTCTTAAGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC	180							
QY	181	TGTTTGGCCACACCGCANACCTGCGCTGACCTCCCATCCCTCTGATCTCTGAGGCTGC	240							
DB	181	TGTTTGGCCACACCGCANACCTGCGCTGACCTCCCATCCCTCTGATCTCTGAGGCTGC	240							
QY	241	CGCTGTGCTCCTGATCCAGCGARGCGCCCATTTGCGGCTCCCAATTTGGGCTAAAGGCTTGC	300							
DB	241	CGCTGTGCTCCTGATCCAGCGARGCGCCCATTTGCGGCTCCCAATTTGGGCTAAAGGCTTGC	300							
QY	301	CATTGTNCTGACGCGCTAAGTGCCTGGGTTGTTTCTAATTGAGCTGAACACTANTCACT	360							
DB	301	CATTGTNCTGACGCGCTAAGTGCCTGGGTTGTTTCTAATTGAGCTGAACACTANTCACT	360							
QY	361	GGGTTCCATGGTTCTCTTCTGTGACCCAGGCTTCTAATAAATCTAATAAATCTAATAAAT	420							
DB	361	GGGTTCCATGGTTCTCTTCTGTGACCCAGGCTTCTAATAAATCTAATAAATCTAATAAAT	420							
QY	421	TGGCCCAAGATTCCCATTCCTTGGAAATCCGTGAGGSCAACGAACTCCAGTCAAGAAATAC	480							
DB	421	TGGCCCAAGATTCCCATTCCTTGGAAATCCGTGAGGSCAACGAACTCCAGTCAAGAAATAC	480							
QY	481	GARGCTTGCCACCAATCTTTGGAAAGCGGCTGTACCTCTTGGAAAGTGGTTCAACCAATC	540							
DB	481	GARGCTTGCCACCAATCTTTGGAAAGCGGCTGTACCTCTTGGAAAGTGGTTCAACCAATC	540							
QY	541	TTGGGAGCTCTGTGAGCAAGGACCCCGGCTGACATTTTGGCRACCMRACGGACATCC	600							
DB	541	TTGGGAGCTCTGTGAGCAAGGACCCCGGCTGACATTTTGGCRACCMRACGGACATCC	600							
QY	601	MAAGTATGGGAAACGTTTCCCGCAAGACAAAAACGCCCTTAAGACGTATTCTGGARAT	660							
DB	601	MAAGTATGGGAAACGTTTCCCGCAAGACAAAAACGCCCTTAAGACGTATTCTGGARAT	660							
QY	661	TGGGAMCAATTGACCTCAGACACTAAGAAAGAAACGACTTATATCTTCTGCAAGTGC	720							
DB	661	TGGGAMCAATTGACCTCAGACACTAAGAAAGAAACGACTTATATCTTCTGCAAGTGC	720							
QY	721	GCCTGGCACTCTGAGGGAAGTATAAATTATAACCACTTTACAGCTAGACTCTTTTG	780							
DB	721	GCCTGGCACTCTGAGGGAAGTATAAATTATAACCACTTTACAGCTAGACTCTTTTG	780							
QY	781	TAGAAAAGGCAATGAGTGAAGTCCCATAGTACAACTTTCTTTTCAATTAAGAGACAA	840							
DB	781	TAGAAAAGGCAATGAGTGAAGTCCCATAGTACAACTTTCTTTTCAATTAAGAGACAA	840							
QY	841	CTCACAAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCTTACAGACTACCTCC	900							
DB	841	CTCACAAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCTTACAGACTACCTCC	900							
QY	901	CTATCCGAGCTCCCGACTCCTTCCCAATTAAGGACCCCTTCAACCCCAATGG	960							
DB	901	CTATCCGAGCTCCCGACTCCTTCCCAATTAAGGACCCCTTCAACCCCAATGG	960							
QY	961	TCCAAAAGGAGATAGACAAAAAGGTAACAGTGAACCAAGAGTGCCAAATATTTCCCAAT	1020							
DB	961	TCCAAAAGGAGATAGACAAAAAGGTAACAGTGAACCAAGAGTGCCAAATATTTCCCAAT	1020							
QY	1021	TATGACCCCTCCCAAGCAGTGGGAGAGAGAAATTCGGCCACGACGAGTGCATGTGCT	1080							
DB	1021	TATGACCCCTCCCAAGCAGTGGGAGAGAGAAATTCGGCCACGACGAGTGCATGTGCT	1080							

QY	1081	TTTTTCTCCAGACTTAAGCAATAAATAACAGACTTTAGTAAATTTCTCAGATAAYCCT	1140
DB	1081		
QY	1141	GATGGCTATATTGRTGTTTACAGGGTTAGGACAAATTTCTTGATCTGACATGGAGAGAT	1200
DB	1141		
QY	1201	ATATATGCTACTGCTAAATCAGACACTAAACCCCAATGAGAGAGTGGCCACCACTAATGCTC	1260
DB	1201		
QY	1261	AGCTGAGRTGTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGATGACAACTA	1320
DB	1261		
QY	1321	GAAGGAAGANAATGATTTCCCAACAGCCAGCARGCAGTTCACAGTCTASACCTCATTTG	1380
DB	1321		
QY	1381	GGGACACAGAAATCAGTAACATGGGAGATTGGTCTGACAGACATTTGCTTAACCTGATGTC	1440
DB	1381		
QY	1441	TASAGGACTAAGGNAACCTASGAGNAARTCTAYGNAATTAATCAATGATGTCACCATTA	1500
DB	1441		
QY	1501	ACACAGGGGAAGGNAAGAAATCTACTGCTTTCTGAGAGACTAAGGAGGCATTTGAG	1560
DB	1501		
QY	1561	GAAGCGTGCCTCTGTGACCTGACTCTTCTGAAGGCCAACTAATCTTAAGCGTAAATTT	1620
DB	1561		
QY	1621	TATCACTCAGTCACTGTCAGACATATTAGAAAAAATTTCAAAGTCTGCGTAGGCCCGGAG	1680
DB	1621		
QY	1681	CAAACTTGAAGACCTTATGAACTTTGCAACCTTCAAAAGTCTGCGTAGGCCCGGAG	1740
DB	1681		
QY	1741	GACAGCGGAAACAGGACAAACGGGATTAATAAAAAAGGCCACCGCTTTAGTCAATGACCTT	1800
DB	1741		
QY	1801	CAGGCAAGTGGACTTTGGAGGCTCTGAAAGGGAAGGCTGGGCAAAATGGAATGCTTAA	1860
DB	1801		
QY	1861	TAGGGCTGCTTCCAGTGGGCTTACAGGACACTTTTAAAGATTTGTCAGATAGAAG	1920
DB	1861		
QY	1921	TAAGCGCCCTTCTGCTCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCC	1980
DB	1921		
QY	1981	CAGGGGACAAAGTCTTTTGAAGTCAAGACCTTAAACCAAGATGATCCAGCAGGAGCTG	2040
DB	1981		
QY	2041	AGGGTGGTGGGGAAGCGCCATCCCATGCCATCACTTCAAGAGGATCACTGGAAGGCCCACTGCCC	2100
DB	2041		
QY	2101	GACCATTTGAGGGCCAGGAGTTGCTCTCTGGACACTGGCGGTCTTCTTAGTCTTACT	2160
DB	2101		

QY	2161	CTTCTGTCGGGCAACTGTCTCTCAGATCTGTCACTATTTCTGAGGGGTCCNTAAGACG	2220
DB	2161		
QY	2221	GGCAGTCACTAGATACCTTTTCCAGCCACTAAGTTTATGAACCTGGGAGCTTTATTTCTTT	2280
DB	2221		
QY	2281	TCACATGCTTTCTAATTAATGCTTGAAGCCCACTACCTTGTGTAGGAGACATTTCTA	2340
DB	2281		
QY	2341	GCAAAAGCAGGGGCCAATTAACCTGAACATAGGAGGAACAACCCGTTTGTGTGNC	2400
DB	2341		
QY	2401	CCTGCTTGAAGGAAGAAATTAATCTGAAGTCTGGGCAACAGAGGACAAATATGACGAGC	2460
DB	2401		
QY	2461	CAAAGAAATGCCGTCCTGTTCAAGTTAACTAAGGATTTCCACTTCTTCCCTACCAA	2520
DB	2461		
QY	2521	GGCAGTACCCCTCAGACCCAGGCCCAACAGAGATTTCAAAGATTTTAAGGACTTAA	2580
DB	2521		
QY	2581	AAGCCCAAGGCTTAGTAAACCATATGATTAACCTCCCTGCTAGTAAATTCGCTAGTGGATGAG	2640
DB	2581		
QY	2641	GAGGCAACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAGAGATCTCAGGATTAATCAATGG	2700
DB	2641		
QY	2701	AGGCCGTTGCTTTTATACCCAGCTGTACTAGCCCTTATATCTGTGMYTTCCCAATATC	2760
DB	2701		
QY	2761	CAGAGGACAGAGTGGTTTACASTCTGGACCTTMAAGATGCTTCTTGATCCCTG	2820
DB	2761		
QY	2821	TACATCCTGACTCTCAATTTCTTTGTTGCTTTTGAAGATCTTCAAAACCCARCATCTCAAC	2880
DB	2821		
QY	2881	TCACCTGGACTTCTTTTACCCCAAGGTTCAAGGATAGYCCCATCTATTTGGCCAGGCAT	2940
DB	2881		
QY	2941	TAGCCCAAGACTTGAGYCARITYMTCACTCTGACACTCTTCTCTGTAGTGTGATG	3000
DB	2941		
QY	3001	ATTTACTTTTTCYGCYVTRTCAGAAAACCTTGTGCCATCAAGCCACCCCAAGRCCTCTTMA	3060
DB	3001		
QY	3061	ATTTACTTTTTCYGCYVTRTCAGAAAACCTTGTGCCATCAAGCCACCCCAAGRCCTCTTMA	3120
DB	3061		
QY	3121	AGGTTAAATACCTTAGGRCCTAARATTTATCAAAAGGACCAARGGCCCTCAGTAGGAAYRVA	3180
DB	3121		
QY	3181	TCCAGCTATCTGGCTTATCTCTCATCYCAAAACCTTAAGCAACTTAGRGRRTTCTTTG	3240
DB	3181		
QY	3241	GCRTAAYAGGYTCTCTGCCAAWATGGATTTCCCAAGTWTGGCRAAATAGCCAGGYCATTTA	3300
DB	3241		



|||||  
3241 GCTAAAGGTTCTGCCGAAWATGATCCCAAGTWTGGCAATATAGCCAGGTCATTA 3300 Db  
|||||  
3301 WATACASTAATPAAGAACTCAGAAAGCCAAATACCCATTTTARTTAAGATGGAYMCTGAA 3360 QY  
|||||  
3301 WATACASTAATPAAGAACTCAGAAAGCCAAATACCCATTTTARTTAAGATGGAYMCTGAA 3360 Db  
|||||  
3361 GYMRAAGTGGCTTTCCAGGCCCTTAAGAAAGCCCTTAACCAAGYCCAGCTGTAAAGVT 3420 QY  
|||||  
3361 GYMRAAGTGGCTTTCCAGGCCCTTAAGAAAGCCCTTAACCAAGYCCAGCTGTAAAGVT 3420 Db  
|||||  
3421 TGCCCAACRGGGCAAGACTTTTSTTYATAYRTACAGAAAGAAACAGRAAYAGCTCTRGGG 3480 QY  
|||||  
3421 TGCCCAACRGGGCAAGACTTTTSTTYATAYRTACAGAAAGAAACAGRAAYAGCTCTRGGG 3480 Db  
|||||  
3481 GTCTTTACACAGRTCCRAGGAYGAGCTTGCAACCCYRTGGCRYACCTGASTAAGGAAAYT 3540 QY  
|||||  
3481 GTCTTTACACAGRTCCRAGGAYGAGCTTGCAACCCYRTGGCRYACCTGASTAAGGAAAYT 3540 Db  
|||||  
3541 GATGTAGTGGCAAGGGTTGRCYTCATTTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTA 3600 QY  
|||||  
3541 GATGTAGTGGCAAGGGTTGRCYTCATTTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTA 3600 Db  
|||||  
3601 GTATCTGAAGCAGTTTAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660 QY  
|||||  
3601 GTATCTGAAGCAGTTTAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660 Db  
|||||  
3661 GTGAAYRGCACTCACTGCTCTAAAGGAGACTTGTGGTGTGACAAACYGTGTTTACTTAAA 3720 QY  
|||||  
3661 GTGAAYRGCACTCACTGCTCTAAAGGAGACTTGTGGTGTGACAAACYGTGTTTACTTAAA 3720 Db  
|||||  
3721 TPTCAGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACTTGTGCACTCTTAAAC 3780 QY  
|||||  
3721 TPTCAGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACTTGTGCACTCTTAAAC 3780 Db  
|||||  
3781 CCAGYCNCACTTCTCCAGCAATGAAGAAAGATARAAYATAAAGTGTCTCAACAARTAAAT 3840 QY  
|||||  
3781 CCAGYCNCACTTCTCCAGCAANTGAAGAAAGATARAAYATAAAGTGTCTCAACAARTAAAT 3840 Db  
|||||  
3841 TCTCAAACTATGCGACCTCGAGGGGACCTTGTAGAGTTTCCYTTGACTGATCCYGACCTT 3900 QY  
|||||  
3841 TCTCAAACTATGCGACCTCGAGGGGACCTTGTAGAGTTTCCYTTGACTGATCCYGACCTT 3900 Db  
|||||  
3901 CAACTGTATATGATGGAGTTCTTTGTAGAAAAGGACTTCGAAAGYGGGGTATGC 3960 QY  
|||||  
3901 CAACTGTATATGATGGAGTTCTTTGTAGAAAAGGACTTCGAAAGYGGGGTATGC 3960 Db  
|||||  
3961 AGTGGTCACTGATTAATGGAATATTTGAAAGTAATCCCTCACTCCAGGAACCTAGTGTCTYA 4020 QY  
|||||  
3961 AGTGGTCACTGATTAATGGAATATTTGAAAGTAATCCCTCACTCCAGGAACCTAGTGTCTYA 4020 Db  
|||||  
4021 GCTRCAGAACTAATAGCCYTCAYTKGGGCACTAGAAATTTAGAGAAAGRAAAAAGGGYAAA 4080 QY  
|||||  
4021 GCTRCAGAACTAATAGCCYTCAYTKGGGCACTAGAAATTTAGAGAAAGRAAAAAGGGYAAA 4080 Db  
|||||  
4081 TATATATACAGACTCTRARTATGCTYACCTAGTONTCCATGCCCCATGMRGCAATATGSR 4140 QY  
|||||  
4081 TATATATACAGACTCTRARTATGCTYACCTAGTONTCCATGCCCCATGMRGCAATATGSR 4140 Db  
|||||  
4141 AGAAAGGGAAATTCCTAACTTCYAGRGAACACCTATCAMACATCAGGAAGCCATTAGGAR 4200 QY  
|||||  
4141 AGAAAGGGAAATTCCTAACTTCYAGRGAACACCTATCAMACATCAGGAAGCCATTAGGAR 4200 Db  
|||||  
4201 ATTATTAATYTGCGWGTACAGAAACCTARAGAGTGGMAGTCTTACACTGCGGGGTATCA 4260 QY  
|||||  
4201 ATTATTAATYTGCGWGTACAGAAACCTARAGAGTGGMAGTCTTACACTGCGGGGTATCA 4260 Db  
|||||  
4261 NAAAGGAAAGRAAGGGAATASAGRGAAYTGCCCAAGCAKATATTGAAGCMWAAAGAGC 4320 QY  
|||||  
4261 NAAAGGAAAGRAAGGGAATASAGRGAAYTGCCCAAGCAKATATTGAAGCMWAAAGAGC 4320 Db  
|||||  
4321 TGCAGGCGAGGACCTCCATTTAGAAATGCTTATTAATCTTCCCTTAGTATAGGTAATCC 4380 QY  
|||||

|||||  
4321 TGCAAGGCAGGACCTCCATTAGAAAATGCTTTATTAACCTTCCCTTAGTATAGGGTAATCC 4380 Db  
|||||  
4381 CTTTCGGGAAACCAAGCCCCAGTAGTACTCAGCAGGAGAAACAGAAATGGGAAACCTTCACGAGG 4440 QY  
|||||  
4381 CTTTCGGGAAACCAAGCCCCAGTAGTACTCAGCAGGAGAAACAGAAATGGGAAACCTTCACGAGG 4440 Db  
|||||  
4441 CAGTTTTCTCCCTCGGGACGGTTAGCCACTGAAAGAGGAGAAATACATTTTTCCTGCTGCAAC 4500 QY  
|||||  
4441 CAGTTTTCTCCCTCGGGACGGTTAGCCACTGAAAGAGGAGAAATACATTTTTCCTGCTGCAAC 4500 Db  
|||||  
4501 TATCAATGGAAATTTACTTAAAAACCTTTTCAATCAATCTTTTCAAACTATCAAGCARAT 4560 QY  
|||||  
4501 TATCAATGGAAATTTACTTAAAAACCTTTTCAATCAATCTTTTCAAACTATCAAGCARAT 4560 Db  
|||||  
4561 CCATCARATGCCCCAATCATTTATTTACTGGAACAGGCCCTTTTCAAACTATCAAGCARAT 4620 QY  
|||||  
4561 CCATCARATGCCCCAATCATTTATTTACTGGAACAGGCCCTTTTCAAACTATCAAGCARAT 4620 Db  
|||||  
4621 AKTCAGGCGCTGTGAATGTGCCARARAAATAATCCCTGCTTATCCGCAAGCTCCTTC 4680 QY  
|||||  
4621 AKTCAGGCGCTGTGAATGTGCCARARAAATAATCCCTGCTTATCCGCAAGCTCCTTC 4680 Db  
|||||  
4681 AGGARAACAAARAAACAGGCCATTTACCTTGRARARAACTGGCAACTGATTTTACCCACAAG 4740 QY  
|||||  
4681 AGGARAACAAARAAACAGGCCATTTACCTTGRARARAACTGGCAACTGATTTTACCCACAAG 4740 Db  
|||||  
4741 CCCAAACCTCAGGAGTTTCAGTATCTACTAGTCTGGGTARATACATTTTACCGGTTGGGCA 4800 QY  
|||||  
4741 CCCAAACCTCAGGAGTTTCAGTATCTACTAGTCTGGGTARATACATTTTACCGGTTGGGCA 4800 Db  
|||||  
4801 RAGGCTTCCCTGTAGGACAGAAAGCCCAAGAGGTAAATAAAGGCACTAGTTTCATGAA 4860 QY  
|||||  
4801 RAGGCTTCCCTGTAGGACAGAAAGCCCAAGAGGTAAATAAAGGCACTAGTTTCATGAA 4860 Db  
|||||  
4861 ATAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGAACAATAGCCCTGCTTTCCAG 4920 QY  
|||||  
4861 ATAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGAACAATAGCCCTGCTTTCCAG 4920 Db  
|||||  
4921 GCCACAGTAAACCCAGGGAGTATCCAGGGCTTGGTATACGATATCACTTACACCTGCGCC 4980 QY  
|||||  
4921 GCCACAGTAAACCCAGGGAGTATCCAGGGCTTGGTATACGATATCACTTACACCTGCGCC 4980 Db  
|||||  
4981 TGAAGGCCCAACAGTCTCAGGGAAGGTGAGAAAAATGAATAAAGCACTCTA 5040 QY  
|||||  
4981 TGAAGGCCCAACAGTCTCAGGGAAGGTGAGAAAAATGAATAAAGCACTCTA 5040 Db  
|||||  
5041 AAAAAGCAAAACCCAGGAAACCCACTCATGCGCTGTCTGTGCTGCTATAGCCTTAAAA 5100 QY  
|||||  
5041 AAAAAGCAAAACCCAGGAAACCCACTCATGCGCTGTCTGTGCTGCTATAGCCTTAAAA 5100 Db  
|||||  
5101 AGAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATAGGAATGCTGTATGGAGG 5160 QY  
|||||  
5101 AGAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATAGGAATGCTGTATGGAGG 5160 Db  
|||||  
5161 CCCTTCATAACCAATGACCTTGTCTGACCCCAAGCAGCCAACTTAGTTTGCAGACATCA 5220 QY  
|||||  
5161 CCCTTCATAACCAATGACCTTGTCTGACCCCAAGCAGCCAACTTAGTTTGCAGACATCA 5220 Db  
|||||  
5221 CCTCCTTAGCCAAATATCAACAAGTTCTTAAAAACATTAAGGAACCTATCCCTGAGAAG 5280 QY  
|||||  
5221 CCTCCTTAGCCAAATATCAACAAGTTCTTAAAAACATTAAGGAACCTATCCCTGAGAAG 5280 Db  
|||||  
5281 AGGAAAAAGAACTATTTCACCCWGTGATGTTAGTCAAGTCCCTTCCTCTTAAT 5340 QY  
|||||  
5281 AGGAAAAAGAACTATTTCACCCWGTGATGTTAGTCAAGTCCCTTCCTCTTAAT 5340 Db  
|||||  
5341 CCCCATCTCTAGATACATCTCTGGGAAGGACCTTACCCAGTCAATTTTATACCCCACTG 5400 QY  
|||||  
5341 CCCCATCTCTAGATACATCTCTGGGAAGGACCTTACCCAGTCAATTTTATACCCCACTG 5400 Db  
|||||  
5401 CGGTTAAAGTGGCTGGAGTCTTGGATACATCACACTTTGAGTCAAAATCCTGGATAC 5460 QY  
|||||  
5401 CGGTTAAAGTGGCTGGAGTCTTGGATACATCACACTTTGAGTCAAAATCCTGGATAC 5460 Db  
|||||



QY	5461	TGCCAAAGGAACCTGAAATCCAGGACAAACCGTAGCTATTCTGTGAACCTCTAGAGG	5520
Db	5461		
QY	5521	ATTTGGCCCTGCTCTTCAAAACAACACGAGGAGAAAGTAACATAAATCAATAATCCCC	5580
Db	5521		
QY	5581	ATGSCCTCCCTTATCATATTTTCTCTKTAAGTTSTTTTACCTSTTTCACTCTCACT	5640
Db	5581		
QY	5641	GCACCCCTCCATGCGCGCTGATGACAGTAGCTCCCTTACCMAGAGTTTCTATGGAGA	5700
Db	5641		
QY	5701	ATGCAGCGTCCCGGAAATATTTGATGCCCCCATCTGTATAGAGTCTTTSTAAGGGAACCCCC	5760
Db	5701		
QY	5761	ACCTTCACTGCCACACCCATATGCCCCCATCTGTATAGAGTCTTTSTAAGGGAACCCCC	5820
Db	5761		
QY	5821	CATGCAAACTCATTTATGGACAGGAAATGATTAATCCTAGTTGTCTGAGGACTT	5880
Db	5821		
QY	5881	GGAGTCACTGTCTGTGGACTTACTTACCACCAACTGGTATGCTGATGGGGTGGAGTT	5940
Db	5881		
QY	5941	CAAGATCAGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCAACTCACCSGGGTACAT	6000
Db	5941		
QY	6001	GGCACTCTAGCCCTTACAAAGACTAGATCTCTCAAAACTAATGAAGACCTCCGTACC	6060
Db	6001		
QY	6061	CATACCTCTAGCCCTTACAAAGACTAGATCTCTCAAACTAATGAAGACCTCCGTACC	6120
Db	6061		
QY	6121	CAAAACCTCTAACTGTGGATATGCTCCCTCGAACTTCAAGCCATATGTTTCAATC	6180
Db	6121		
QY	6181	CCTGTACCTGAAACAATGGAACTTCAAGCAAGAAATAAACACCACTTCCGTTTATGA	6240
Db	6181		
QY	6241	GGACCTCTGTTTCCAAATSTGGAATAACCCATACCTCAAACTCACTGTGTAAATTT	6300
Db	6241		
QY	6301	AGCAATACATACACAACTCCCAATGATCAGGTGGGTAACTCTCTCCACACAA	6360
Db	6301		
QY	6361	ATAGTCTGCTACCCCTCAGGAATATTTTGTCTGTGGTACCTCAGCCATCTGTTTGG	6420
Db	6361		
QY	6421	AATGGCTCTCAAGATCTATGCTGCTCTCATCTTATGTCGCCCCCATGRCATCTAC	6480
Db	6421		
QY	6481	ACTGAAACAAGATTTATACAGTTATGTCATATCTAAGCCCGCAACAAAGAGTACCATT	6540
Db	6481		

QY	6541	CTTCTCTTTTGTATAGGACGAGGAGTCTAGGTGCACTAGGTACTGGCAATTGGCGGTATC	6600
Db	6541		
QY	6601	ACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG	6660
Db	6601		
QY	6661	GTGCGCGACTCCCTGTGTCACTTGTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTT	6720
Db	6661		
QY	6721	CRAAATCGAGAGCTTTAGACTTGTAAACCGCTGAPAGAGGGGAACTGTATTTTAA	6780
Db	6721		
QY	6781	GGGGAAGATGCTGTTTATTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAGAA	6840
Db	6781		
QY	6841	ATTCSAGATCGAATACAACGTAKAGCAGAGAGCTTCGAAACAACCTGGGCGCTC	6900
Db	6841		
QY	6901	CTCAGCCRAATGAGTGGCTGGATTCTCCCTCTTATAGGACCTCTAGCAGCTATAATTTG	6960
Db	6901		
QY	6961	CTACTCTCTTTGGACCTGTATCTTTTACCTCTCTTGTAACTTTCTCTCTCCAGATC	7020
Db	6961		
QY	7021	GAAGCTGTAAACTCAAAATGGAGCCCAAGATGCACTCAAGACTAAGATCTTACCAGCA	7080
Db	7021		
QY	7081	CCCTGGACCGGCTGTATAGCCCAAGATCTGTATGTTTAAATGACATCAAGGACACCTCTCT	7140
Db	7081		
QY	7141	GAGGAATCTCAGCTGCACAACTCTACTACGCCCAATTTACGAGGAGGAGTGTAGGC	7200
Db	7141		
QY	7201	GGTGTGGCCCACTCCCAACAGCACTTAGGTTTCTGTGAGATGGGGACCTGAGA	7260
Db	7201		
QY	7261	GACGAGCTAGCTGGATTCTTAGGCTGATTAAGAAATCCVTAAGCCTAGSTGGGAAGGTG	7320
Db	7261		
QY	7321	ACCAATCCACCTTTAAACACGGGCTTGCACCTTAGYTACACCTGACCAATCAGAGAG	7380
Db	7321		
QY	7381	CTCACTAAAATGCTAATTAAGGCAAGACAGGAGGTAAAGAAATAGCCAAATCATTTATTC	7440
Db	7381		
QY	7441	MTGAGAGCACAGCAGGAGGACAAATGATCGGGATATAAACCCCAAGTCTCGAGCCCGCAA	7500
Db	7441		
QY	7501	CGGCAACCCCTTTGGTCCCTCCCTTTGTATGGAGCTCTGTTTCAATGCTATTTTAC	7560
Db	7501		
QY	7561	TCTATTAAATCTTGACCTGCR	7582
Db	7561		

RESULT 3  
US-10-450-763-4430  
; Sequence 4430, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; PRIOR FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 4430  
; LENGTH: 7466  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIMILAR  
; LOCATION: (1378)..(2271)  
; OTHER INFORMATION: 84% homologous to multiple sclerosis associated retrovirus  
; OTHER INFORMATION: polyprotein, accession number AF009668, Smith-Waterman Score=1279.  
US-10-450-763-4430

Query Match 12.9%; Score 979; DB 10; Length 7466;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1269; Conservative 0; Mismatches 1; Indels 2; Gaps 2;  
  
QY 1476 GAATTACTCAATGATGTCACCATACACAGGGGAGGAGAAATCTACTGCTTTC 1535  
DB 5407 GAATTACTCAATGATGTCACCATACACAGGGGAGGAGAAATCTACTGCTTTC 5466  
  
QY 1536 TGGAGAGACTAAGGAGGAGGATGAGGAGCGTCCCTCTGTCACCTGACTCTTCTGAG 1595  
DB 5467 TGGAGAGACTAAGGAGGAGGATGAGGAGCGTCCCTCTGTCACCTGACTCTTCTGAG 5526  
  
QY 1596 GCCAATCAATCTTAAAGCGTAGTTTATCACTCAGTCAGTCAGACATTAGAAAAAATCT 1655  
DB 5527 GCCAATCAATCTTAAAGCGTAGTTTATCACTCAGTCAGTCAGACATTAGAAAAAATCT 5586  
  
QY 1656 TCAAAAGTCTGCCGTAGGCGCGAGCAAAATCTTAGAAACCTTATTGAACCTTGGCAACATC 1715  
DB 5587 TCAAAAGTCTGCCGTAGGCGCGAGCAAAATCTTAGAAACCTTATTGAACCTTGGCAACATC 5646  
  
QY 1716 GGTTTTATATATAGAGATCAGGAGGAGCGGAGCAAGGAGCAAAACCGGATTAAGAAA 1775  
DB 5647 GGTTTTATATATAGAGATCAGGAGGAGCGGAGCAAGGAGCAAAACCGGATTAAGAAA 5706  
  
QY 1776 AGSCCAACCGCTTTAGTCATGACCTCAGGCAAGTGAGCTTTGGAGCTCTGCAAGAGGCA 1835  
DB 5707 AGSCCAACCGCTTTAGTCATGACCTCAGGCAAGTGAGCTTTGGAGCTCTGCAAGAGGCA 5766  
  
QY 1836 AAGAGCTGGGCAAAATGAAATGCTTAATAGGCTTGTCTTCAGTGGCGGTCTTACAGGACAT 1895  
DB 5767 AAGAGCTGGGCAAAATGAAATGCTTAATAGGCTTGTCTTCAGTGGCGGTCTTACAGGACAT 5826  
  
QY 1896 TTAAGAGATGTCAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAG 1955  
DB 5827 TTAAGAGATGTCAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAG 5886  
  
QY 1956 GGAATCAGCTGGAGGCGGCAATGCGCCAGGAGCAAAAGGTCTTTTGTAGTCAGAGGCACTA 2015  
DB 5887 GGAATCAGCTGGAGGCGGCAATGCGCCAGGAGCAAAAGGTCTTTTGTAGTCAGAGGCACTA 5946  
  
QY 2016 ACCAGATGATCCAGCAGCAGGAGCTAGAGGTGCTGGGCAAGCGCCATCCCATGCGCATCA 2075  
DB 5947 ACCAGATGATCCAGCAGCAGGAGCTAGAGGTGCTGGGCAAGCGCCATCCCATGCGCATCA 6006

QY 2076 CCCTCACAGAGCCCTGGGTATGCTTGACCATTTAGGGGCGAGGAAGGTGTCTCTCGGACA 2135  
DB 6007 CCCTCACAGAGCCCTGGGTATGCTTGACCATTTAGGGGCGAGGAAGGTGTCTCTCGGACA 6066  
  
QY 2136 CTGGTGGCGTCTTCTTACTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2195  
DB 6067 CTGGTGGCGTCTTCTTACTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6126  
  
QY 2196 CTATTCTCAGGGGCTCCTTAAGACGGGCGAGTCAGTACTCTTTTCCAGGACCTAAGT 2255  
DB 6127 CTATTCTCAGGGGCTCCTTAAAGCGGCGAGTCAGTACTCTTTTCCAGGACCTAAGT 6185  
  
QY 2256 TATGAACCTGGGAGCTTTTATTCTTTTACATGCTTTTCTAAATATGCTTGAAGCCCCAC 2315  
DB 6186 TATGAACCTGGGAGCTTTTATTCTTTTACATGCTTTTCTAAATATGCTTGAAGCCCCAC 6245  
  
QY 2316 TACCTTGTAGGAGAGACATTTAGCAAAAAGCAGGGGCCATTTATACCTGGAACATAGG 2375  
DB 6246 TACCTTGTAGGAGAGACATTTAGCAAAAAGCAGGGGCCATTTATACCTGGAACATAGG 6305  
  
QY 2376 AGAGGAAACACCGTTTGTGTCCTGCTTGGAGGAAGGAATTAATCTGAGTCTGGG 2435  
DB 6306 AGAGGAAACACCGTTTGTGTGTCCTGCTTGGAGGAAGGAATTAATCTGAGTCTGGG 6364  
  
QY 2436 CAACAGAGGACAATATGAGCAGGCAAAAGATGCGCTCTGTTTCAAGTTTAAACTAAAG 2495  
DB 6365 CAACAGAGGACAATATGAGCAGGCAAAAGATGCGCTCTGTTTCAAGTTTAAACTAAAG 6424  
  
QY 2496 GATTCACCTCTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2555  
DB 6425 GATTCACCTCTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6484  
  
QY 2556 TTCCAAAAGATTTGTTAAGGACTTAAAGCCCAAGGCTTAGTAAAAACCATGCACTCC 2615  
DB 6485 TTCCAAAAGATTTGTTAAGGACTTAAAGCCCAAGGCTTAGTAAAAACCATGCACTCC 6544  
  
QY 2616 TGCAGTAATTCGCTAGTGGATTGAGGAGGACAGAAACCCAGTGACAGTGGAGGTTAG 2675  
DB 6545 TGCAGTAATTCGCTAGTGGATTGAGGAGGACAGAAACCCAGTGACAGTGGAGGTTAG 6604  
  
QY 2676 TGCAGTAATTCGCTAGTGGATTGAGGAGGACAGAAACCCAGTGACAGTGGAGGTTAG 2735  
DB 6605 TGCAGTAATTCGCTAGTGGATTGAGGAGGACAGAAACCCAGTGACAGTGGAGGTTAG 6664  
  
QY 2736 CCTTATCTGTG 2747  
DB 6665 CCTTATCTGTG 6676

RESULT 4  
US-10-632-793-29  
; Sequence 29, Application US/10632793  
; Publication No. US20040048298A1  
; GENERAL INFORMATION:  
; APPLICANT: PARANHOS-BACCALA, Glaucia  
; APPLICANT: MALLET, Francois  
; APPLICANT: VOISSET, Cecile  
; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN  
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT  
; FILE REFERENCE: 110048  
; CURRENT APPLICATION NUMBER: US/10/632,793  
; CURRENT FILING DATE: 2003-08-04  
; PRIOR APPLICATION NUMBER: US/09/869,927  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: PCT/FR00/00144  
; PRIOR FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: FR 99/00888  
; PRIOR FILING DATE: 1999-01-21  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 29  
; LENGTH: 2372  
; TYPE: DNA

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1191)..(1191)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1213)..(1213)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2089)..(2089)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2274)..(2274)
; OTHER INFORMATION: n = a or g or c or t/u
US-10-632-793-29

Query Match      11.7%; Score 888; DB 8; Length 2372;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 988; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1507 GGGAGGGAGAAATCCCTACTGCTTCTGGAGAGACTAAGGAGGCATTGAGGAAGCG 1566
DB 1383 GGGAGGGAGAAATCCCTACTGCTTCTGGAGAGACTAAGGAGGCATTGAGGAAGCG 1442

QY 1567 TGCCTCTGTCACTGACTTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTATCAC 1626
DB 1443 TGCCTCTGTCACTGACTTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTTATCAC 1502

QY 1627 TCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTTCGCCGTAGGCCGAGCAAAAC 1686
DB 1503 TCAGTCAGCTGCAGACATTAGAAAAAACTTCAAAAGTTCGCCGTAGGCCGAGCAAAAC 1562

QY 1687 TTAGAAACCTCTTGAACCTTGCAACVTCGGTTTTTTTAAATAGAGATCAGGAGGAGCAG 1746
DB 1563 TTAGAAACCTCTTGAACCTTGCAACVTCGGTTTTTTTAAATAGAGATCAGGAGGAGCAG 1622

QY 1747 GCGGAACAGGACAAACCGGATTAATAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1806
DB 1623 GCGGAACAGGACAAACCGGATTAATAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1682

QY 1807 AGTGGACTTTGGAGGCTCTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1866
DB 1683 AGTGGACTTTGGAGGCTCTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1742

QY 1867 TTGCTTCCAGTGGGTCTACAGGACACTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1926
DB 1743 TTGCTTCCAGTGGGTCTACAGGACACTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1802

QY 1927 GCGCCTTCCTGATGCGCTTATTTCAAGGGAATCACTGGAAGGCGCCACTGCGCCAGGGG 1986
DB 1803 GCGCCTTCCTGATGCGCTTATTTCAAGGGAATCACTGGAAGGCGCCACTGCGCCAGGGG 1862

QY 1987 ACAAGGCTTTTGAAGTCAAGAGCCACTAACAGATGATCCAGCAGCAGGAGCTGAGGGTG 2046
DB 1863 ACAAGGCTTTTGAAGTCAAGAGCCACTAACAGATGATCCAGCAGCAGGAGCTGAGGGTG 1922

QY 2047 CCTGGGGCAGGCCCATCCATGCCATCAACCTCAAGAGCCCTGGGTATGCTTGACCAT 2106
DB 1923 CCTGGGGCAGGCCCATCCATGCCATCAACCTCAAGAGCCCTGGGTATGCTTGACCAT 1982

QY 2107 TGAGGCGCCAGGAAGGTGTCTCTCGACACTGGTGGGTCTTCTTAGTCTTACTCTTCG 2166
DB 1983 TGAGGCGCCAGGAAGGTGTCTCTCGACACTGGTGGGTCTTCTTAGTCTTACTCTTCG 2042

QY 2167 TCCCGGACAACTGTCTCTCAGATCTCTCAATTTCTGAGGGGTCCTTAAGAGCGGCAGT 2226
DB 2043 TCCCGGACAACTGTCTCTCAGATCTCTCAATTTCTGAGGGGTCCTTAAGAGCGGCAGT 2102

QY 2227 CACTAGATATTTTCCAGGCACTAAGTTATGAAGTGGGAGCTTTTATCTTTTTCACAT 2286
DB 1383 GGGAGGGAGAAATTCCTACTGCTTCTGAGAGACTAAGGAGGCATTGAGGAAGCG 1566
DB 1383 GGGAGGGAGAAATTCCTACTGCTTCTGAGAGACTAAGGAGGCATTGAGGAAGCG 1442
```

```

DB 2103 CACTAGATATTTTCCAGGCCACTAAGTTATGAAGTGGGAGCTTTATTTCTTTTTCACAT 2162
QY 2287 GCTTTTCTAATTAATGCTTGAAGCCCACTTACCTTGTAGGGAGAGACATTTCTAGCAAAA 2346
DB 2163 GCTTTTCTAATTAATGCTTGAAGCCCACTTACCTTGTAGGGAGAGACATTTCTAGCAAAA 2222
QY 2347 GCAGGGGCCATTATACACCTGAACATAGGAGAGGAACACCCGTTTGTGTGTCCTGCT 2406
DB 2223 GCAGGGGCCATTATACACCTGAACATAGGAGAGGAACACCCGTTTGTGTGTCCTGCT 2282
QY 2407 TGAGGAAGGAATTAATCTTGAAGTCTGGGCAACAGAGAGGACATATATGGAGGCAAAAGA 2466
DB 2283 TGAGGAAGGAATTAATCTTGAAGTCTGGGCAACAGAGAGGACATATATGGAGGCAAAAGA 2342
QY 2467 ATGCCCGTCTCTGTTCAAGTTAACTAAAGG 2496
DB 2343 ATGCCCGTCTCTGTTCAAGTTAACTAAAGG 2372

RESULT 5
US-10-717-580-10
; Sequence 10, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
; APPLICANT: BESEME, Frederic
; APPLICANT: BLOND, Jean-Luc
; APPLICANT: BOUTON, Olivier
; APPLICANT: MANDRAND, Bernard
; APPLICANT: MALLET, Francois
; APPLICANT: PERRON, Hervé
; TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DIS
; FILE REFERENCE: 105045
; CURRENT APPLICATION NUMBER: US/10/717,580
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US/09/446,024A
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/FR98/01442
; PRIOR FILING DATE: 1998-07-06
; PRIOR APPLICATION NUMBER: FR 97/08815
; PRIOR FILING DATE: 1997-07-07
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 10
; LENGTH: 2372
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1191)..(1191)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1213)..(1213)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2089)..(2089)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2274)..(2274)
; OTHER INFORMATION: n = any nucleotide
US-10-717-580-10

Query Match      11.7%; Score 888; DB 9; Length 2372;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 988; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1507 GGGAGGGAGAAATTCCTACTGCTTCTGAGAGACTAAGGAGGCATTGAGGAAGCG 1566
DB 1383 GGGAGGGAGAAATTCCTACTGCTTCTGAGAGACTAAGGAGGCATTGAGGAAGCG 1442
```

```

QY 1567 TGCCTCTCTGTCCACCTGACTCTTCTGAAGGCCAACTAACTTTAAAGCGTAAGTTTATCAC 1626
Db 1443 TGCCTCTCTGTCCACCTGACTCTTCTGAAGGCCAACTAACTTTAAAGCGTAAGTTTATCAC 1502
QY 1627 TCAGTCAGTCGAGACATTTAGAAAAAATCTTCAAAAGTCTGCGTAGGCGCCGGAGCAAAAC 1686
Db 1503 TCAGTCAGTCGAGACATTTAGAAAAAATCTTCAAAAGTCTGCGTAGGCGCCGGAGCAAAAC 1562
QY 1687 TTAGAAACCCATTATGAATCTTGCAATCTGCGTTTATTAATATAGAGATCAGAGAGAGAG 1746
Db 1563 TTAGAAACCCATTATGAATCTTGCAATCTGCGTTTATTAATATAGAGATCAGAGAGAGAG 1622
QY 1747 GCGGAACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCCCTCAGGCA 1806
Db 1623 GCGGAACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCCCTCAGGCA 1682
QY 1807 AGTGACATTTTGAGGCTCTGGAAGAGGGAAGAGCTGGGCAAAATTTGAATGCGCTTAATAGGCG 1866
Db 1683 AGTGACATTTTGAGGCTCTGGAAGAGGGAAGAGCTGGGCAAAATTTGAATGCGCTTAATAGGCG 1742
QY 1867 TTGCTTCCAGTCGGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCC 1926
Db 1743 TTGCTTCCAGTCGGGTCTACAAGGACACTTTAAAAAAGATTGTCCAAGTAGAAGTAAGCC 1802
QY 1927 GCCCCTTCGTCCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGG 1986
Db 1803 GCCCCTTCGTCCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGG 1862
QY 1987 ACAAAAGGCTTTTGAAGTCAGAAAGCCAACTAACACAGATGATCCAGCAGCAGGACTGAGGGTG 2046
Db 1863 ACAAAAGGCTTTTGAAGTCAGAAAGCCAACTAACACAGATGATCCAGCAGCAGGACTGAGGGTG 1922
QY 2047 CTTGGGGCAAGCCCAATGCCATGCCATCAACCCCTCAAGAGCCCTGGGTATGCTTGACCAT 2106
Db 1923 CTTGGGGCAAGCCCAATGCCATGCCATCAACCCCTCAAGAGCCCTGGGTATGCTTGACCAT 1982
QY 2107 TCAGGGCCAGGAAGGTGTCTCTCGGACACTGGTGGGTCTCTTAGTCTTACTCTTCTG 2166
Db 1983 TCAGGGCCAGGAAGGTGTCTCTCGGACACTGGTGGGTCTCTTAGTCTTACTCTTCTG 2042
QY 2167 TCCCGGACAACTGTCTCTCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCGAT 2226
Db 2043 TCCCGGACAACTGTCTCTCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCGAT 2102
QY 2227 CACTAGATATCTTTCAGGACCACTAAAGTTATGAACCTGGGAGCTTTATTTTCAAT 2286
Db 2103 CACTAGATATCTTTCAGGACCACTAAAGTTATGAACCTGGGAGCTTTATTTTCAAT 2162
QY 2287 GCTTTTCTAATTTATGCTTGAAGCCCACTACCTTGTGAGGAGAGACATTTCTAGCAAA 2346
Db 2163 GCTTTTCTAATTTATGCTTGAAGCCCACTACCTTGTGAGGAGAGACATTTCTAGCAAA 2222
QY 2347 GCAGGGGCCATTATACACCTGAACATAGGAGAGGAACACCCGTTTGTGTCCTGCT 2406
Db 2223 GCAGGGGCCATTATACACCTGAACATAGGAGAGGAACACCCGTTTGTGTCCTGCT 2282
QY 2407 TGAGGAAGGAATTAATCTTGAAGTCTGGGCAACAGAGGACAAATATGGAAGGCAAAAGA 2466
Db 2283 TGAGGAAGGAATTAATCTTGAAGTCTGGGCAACAGAGGACAAATATGGAAGGCAAAAGA 2342
QY 2467 ATGCCCGTCTGTTCAAGTTAACTAAAGG 2496
Db 2343 ATGCCCGTCTGTTCAAGTTAACTAAAGG 2372

```

```

RESULT 6
US-11-028-539-1
; Sequence 1, Application US/11028539
; Publication No. US20050118573A1
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois

```

```

; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
; TITLE OF INVENTION: ENDGENOUS RETROVIRAL MOTIFS
; FILE REFERENCE: 200936USPCT
; CURRENT APPLICATION NUMBER: US/11/028.539
; CURRENT FILING DATE: 2005-01-05
; PRIOR APPLICATION NUMBER: US/09/719,554
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2599
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-028-539-1

Query Match      10.3%; Score 779; DB 13; Length 2599;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 4802 AGGCTTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTTAATAAGSCACTAGTTTCATGAAA 4861
Db 151 AGGCTTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTTAATAAGSCACTAGTTTCATGAAA 210
QY 4862 TAATTCACAGATTCCGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTCTTCCAGG 4921
Db 211 TAATTCACAGATTCCGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTCTTCCAGG 270
QY 4922 CCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTCGCGCT 4981
Db 271 CCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTCGCGCT 330
QY 4982 GAAGGCCACAGTCTCTAGGGAAGGTCGAGAAAAATGAATGAAYACTCAAAAGACATCTAA 5041
Db 331 GAAGGCCACAGTCTCTAGGGAAGGTCGAGAAAAATGAATGAAYACTCAAAAGACATCTAA 390
QY 5042 AAAAGCAAAACCCAGGAACCCACCTCATCGCCCTGTCTGTGCTATAGCTTTAAAAA 5101
Db 391 AAAAGCAAAACCCAGGAACCCACCTCATCGCCCTGTCTGTGCTATAGCTTTAAAAA 450
QY 5102 GAATCTGCAACTTTTCCCAAAAGCAGGACTTAGCCCATAGCAATGCTGTATGGAAGGC 5161
Db 451 GAATCTGCAACTTTTCCCAAAAGCAGGACTTAGCCCATAGCAATGCTGTATGGAAGGC 510
QY 5162 CTTTCATTAACCAATGACCTTGTGTGACCCCAAGACAGCCAACTTAGTTGCAAGATCAC 5221
Db 511 CTTTCATTAACCAATGACCTTGTGTGACCCCAAGACAGCCAACTTAGTTGCAAGATCAC 570
QY 5222 CTCCTTAGCCAAATATCAACAAGTTCTTAAAAACATTACAAGGAACCTATCCCTGAGAGA 5281
Db 571 CTCCTTAGCCAAATATCAACAAGTTCTTAAAAACATTACAAGGAACCTATCCCTGAGAGA 630
QY 5282 GGGAAAGAACTATTTCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAAATTC 5341
Db 631 GGGAAAGAACTATTTCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAAATTC 690
QY 5342 CCCATCCCTAGATACATCTCTGGGAGGACCCCTACCCAGTCAATTTTATYACCCCACTGC 5401
Db 691 CCCATCCCTAGATACATCTCTGGGAGGACCCCTACCCAGTCAATTTTATCTACCCCACTGC 750
QY 5402 GGTTHAAGTGGCTGGAGTGGAGTCTTGGATATACATCAGCTTGAGTCAAACTCTGGATCT 5461
Db 751 GGTTHAAGTGGCTGGAGTGGAGTCTTGGATATACATCAGCTTGAGTCAAACTCTGGATCT 810
QY 5462 GCCAAAGGAACCTGAAATCCAGGAGACAAAGCTAGCTATTCTCTGTAACCTCTAGAGA 5521
Db 811 GCCAAAGGAACCTGAAATCCAGGAGACAAAGCTAGCTATTCTCTGTAACCTCTAGAGA 870
QY 5522 TTTGCGCTGCTCTTCAAAACAAACCCAGGAGGAAAGTAACCTTAAATCATATAATCCCCA 5581
Db 871 TTTGCGCTGCTCTTCAAAACAAACCCAGGAGGAAAGTAACCTTAAATCATATAATCCCCA 929

```

QY 5582 TGGSCCTCCCTATCATATTTTCTCTKTAAGTGTSTTTTACCCCTSTTTTCACTCTCACTG 5641  
 Db 930 TGGCCCTCCCTATCATATTTTCTCTTACTGTCTTTTACCCCTCTTTTCACTCTCACTG 989  
 QY 5642 CACCCCTCCATGCGCTGTATGACCAAGTAGTCCCTCTYACMAGAGTTTCTATGGAGAA 5701  
 Db 990 CACCCCTCCATGCGCTGTATGACCAAGTAGTCCCTCTYACMAGAGTTTCTATGGAGAA 1049  
 QY 5702 TGCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCA 5761  
 Db 1050 TGCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCA 1109  
 QY 5762 CTTTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCAATGC 5821  
 Db 1110 CTTTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCAATGC 1169  
 QY 5822 ATGCAAAATCTCAATATTGACAGGAGAAATGATTAATCTCTAGTGTCTCTGGAGGACTTG 5881  
 Db 1170 ATGCAAAATCTCAATATTGACAGGAGAAATGATTAATCTCTAGTGTCTCTGGAGGACTTG 1229  
 QY 5882 GAGTCACTGTCTGTGGACTTACTTACCCAAACTGCTATGCTGTATGCTGGGGGTGGAGTTC 5941  
 Db 1230 GAGTCACTGTCTGTGGACTTACTTACCCAAACTGCTATGCTGTATGCTGGGGGTGGAGTTC 1289  
 QY 5942 AAGATCAGGCAAGAGAAATCATGTAAAGAAAGTAATCTCCCAACTCACCGSGGTACATG 6001  
 Db 1290 AAGATCAGGCAAGAGAAATCATGTAAAGAAAGTAATCTCCCAACTCACCGSGGTACATG 1349  
 QY 6002 GCACCTCTAGCCCTTACAAGGACTAGATCTCTCAAAACTTACATGAACCCCTCCGTACCC 6061  
 Db 1350 GCACCTCTAGCCCTTACAAGGACTAGATCTCTCAAAACTTACATGAACCCCTCCGTACCC 1409  
 QY 6062 ATACTCGCTGTGAAGCTATTATTAATACACCTCTACTGGGCTCCATGAGGTCTCGGCC 6121  
 Db 1410 ATACTCGCTGTGAAGCTATTATTAATACACCTCTACTGGGCTCCATGAGGTCTCGGCC 1469  
 QY 6122 AAAACCTCTACTACTGTGTGATATGCTCCCTGGAATCTTCAAGCCATATGTTTCAATCC 6181  
 Db 1470 AAAACCTCTACTACTGTGTGATATGCTCCCTGGAATCTTCAAGCCATATGTTTCAATCC 1529  
 QY 6182 CTGTACTGAACCAATGGAACAACTTCAGCACAGAAATAAACAACACTTCGGTTTATAGTAG 6241  
 Db 1530 CTGTACTGAACCAATGGAACAACTTCAGCACAGAAATAAACAACACTTCGGTTTATAGTAG 1589  
 QY 6242 GACCTCTGTTTCCAAATSTGGAATAAACCATACCTCAAACCTCACTGTGTAAATTTA 6301  
 Db 1590 GACCTCTGTTTCCAAATSTGGAATAAACCATACCTCAAACCTCACTGTGTAAATTTA 1649  
 QY 6302 GCAATACTACATACAAACCACTCCCAATGATCAGGTGGGTAACTCTCCACACAA 6361  
 Db 1650 GCAATACTACATACAAACCACTCCCAATGATCAGGTGGGTAACTCTCCACACAA 1709  
 QY 6362 TAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCGTTGTTTGA 6421  
 Db 1710 TAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCGTTGTTTGA 1769  
 QY 6422 ATGGCTCTTCAGAAATCTATGTCTCTCTCAATCTTATAGTCCCTCCATGRCATCTACA 6481  
 Db 1770 ATGGCTCTTCAGAAATCTATGTCTCTCTCAATCTTATAGTCCCTCCATGRCATCTACA 1829  
 QY 6482 CTGAACAAGATTTATACAGTTATGTCAATCTTAAGCCCGCAACAAAGAGTACCCATTC 6541  
 Db 1830 CTGAACAAGATTTATACAGTTATGTCAATCTTAAGCCCGCAACAAAGAGTACCCATTC 1889  
 QY 6542 TTCTTTTGTATAGAGCAGAGGTCTAGGTGCACTGCTGCTATGCGATGCGGTATCA 6601  
 Db 1890 TTCTTTTGTATAGAGCAGAGGTCTAGGTGCACTGCTGCTATGCGATGCGGTATCA 1949  
 QY 6602 CAACTCTACTCAGTTCTTACTACAAACTATCTCAAGAACTAAATGGGGGACATGGACGGG 6661  
 Db 1950 CAACTCTACTCAGTTCTTACTACAAACTATCTCAAGAACTAAATGGGGGACATGGACGGG 2009  
 QY 6662 TCCCGACTCCCTGGTCACTTGGCAAGATCAACTTAACTCCCTCCTAGCAGCAGTAGTCTTTC 6721

Db 2010 TCGCCGACTCCCTGGTCACTTGAAGATCACTTAACTCCCTAGCAGCAGTAGTCTCCTC 2069  
 RESULT 7  
 US-09-854-867-21  
 ; Sequence 21, Application US/09854867  
 ; Publication No. US20030224356A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JOAN, KNOELL H  
 ; APPLICANT: ROGAN, PETER K  
 ; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATION  
 ; FILE REFERENCE: 30307  
 ; CURRENT APPLICATION NUMBER: US/09/854,867  
 ; CURRENT FILING DATE: 2003-05-08  
 ; NUMBER OF SEQ ID NOS: 613  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 21  
 ; LENGTH: 8523  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: repeat region  
 ; LOCATION: (1)..(8523)  
 ; OTHER INFORMATION: herv17  
 ; US-09-854-867-21  
 Query Match 10.3%; Score 779; DB 3; Length 8523;  
 Best Local Similarity 98.9%; Pred. No. 0;  
 Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;  
 QY 4802 AGGCTTCCCTGTAGGACAGAAAAAGGCCCAAGAGTAATAAAGGCACCTAGTTTCATGAAA 4861  
 Db 6071 AGGCTTCCCTGTAGGACAGAAAAAGGCCCAAGAGTAATAAAGGCACCTAGTTTCATGAAA 6130  
 QY 4862 TAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGCACATAGCCCTGCTTTCAGG 4921  
 Db 6131 TAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGCACATAGCCCTGCTTTCAGG 6190  
 QY 4922 CCAAGTAAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACATCTGCGCCCT 4981  
 Db 6191 CCAAGTAAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACATCTGCGCCCT 6250  
 QY 4982 GAAGGCCACAGTCTCAGGGAAGGTCGAGAAAAATGAATGAAAYACTCAAAGGACATCTAA 5041  
 Db 6251 GAAGGCCACAGTCTCAGGGAAGGTCGAGAAAAATGAATGAAAYACTCAAAGGACATCTAA 6310  
 QY 5042 AAAAGCAACCCAGGAAACCCACCTCAATGGCGCTGTCCTGCTATAGCCTTAAAAA 5101  
 Db 6311 AAAAGCAACCCAGGAAACCCACCTCAATGGCGCTGTCCTGCTATAGCCTTAAAAA 6370  
 QY 5102 GAATCTGCAACTTTCCTCAAAAGCAGGACTTACCCCATACGAAATGCTGTATGGAAGGC 5161  
 Db 6371 GAATCTGCAACTTTCCTCAAAAGCAGGACTTACCCCATACGAAATGCTGTATGGAAGGC 6430  
 QY 5162 CTTTCATAACCAATGACCTTGTGCTTACCCAGGAGGACCACTTAGTTGACAGACATCAC 5221  
 Db 6431 CTTTCATAACCAATGACCTTGTGCTTACCCAGGAGGACCACTTAGTTGACAGACATCAC 6490  
 QY 5222 CTCCTTAGCCAAATATCAAGTTCTTAAAAATTAAGGAACTATCCCTGAGAGA 5281  
 Db 6491 CTCCTTAGCCAAATATCAAGTTCTTAAAAATTAAGGAACTATCCCTGAGAGA 6550  
 QY 5282 GGGAAAGAACTATTCCACCCWGTGACATGGTATAGTCAAGTCCCTTCCTCTAATTC 5341  
 Db 6551 GGGAAAGAACTATTCCACCCWGTGACATGGTATAGTCAAGTCCCTTCCTCTAATTC 6610  
 QY 5342 CCCATCCCTAGATACATCTGCGGAGGACCCCTACCCAGTCAATTTATYACCCCACTCC 5401  
 Db 6611 CCCATCCCTAGATACATCTGCGGAGGACCCCTACCCAGTCAATTTATCTACCCCACTCC 6670  
 QY 5402 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGGAGTCAATCTCGGATCT 5461

Db 6671 GGTAAAGTGGCTGGAGTCTTGGATACATCACACTTGAATCAATCTCGATACT 6730  
 Qy 5462 GCCAAGGAACTGAAATCCAGGAGACACGCTAGCTATTCTGTGAACCTCTAGAGGA 5521  
 Db 6731 GCCAAGGAACTGAAATCCAGGAGACACGCTAGCTATTCTGTGAACCTCTAGAGGA 6790  
 Qy 5522 TTTCGGCTGCTCTTCAACAACACACGAGGAGAAAGTAACATAAATCATAAATCCCCA 5581  
 Db 6791 TTTCGGCTGCTCTTCAACAACACACGAGGAGAAAGTAACATAAATCATAAATCCCCA 6849  
 Qy 5582 TGGSCCTCCCTTATCATATTTTCTCTPASTGTTSTTTTACCTSTTTTCACTCTCACTG 5641  
 Db 6850 TGGSCCTCCCTTATCATATTTTCTCTTACTGTTCTTTTACCTCTTCTCACTCTCACTG 6909  
 Qy 5642 GACCCCTCCATGCGCTGTATGACAGTAGTCTCCCTTACCMAGAGTTTCTATGGAGAA 5701  
 Db 6910 GACCCCTCCATGCGCTGTATGACAGTAGTCTCCCTTACCAAGAGTTTCTATGGAGAA 6969  
 Qy 5702 TGCAGGCTCCGGAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCA 5761  
 Db 6970 TGCAGGCTCCGGAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCA 7029  
 Qy 5762 CTTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTTCATGC 5821  
 Db 7030 CTTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTTCATGC 7089  
 Qy 5822 ATGCAAACTCATTTATTGGACAGGAAATGATTAATCTTCTAGTTGCTCTGGAGACTTG 5881  
 Db 7090 ATGCAAACTCATTTATTGGACAGGAAATGATTAATCTTCTAGTTGCTCTGGAGACTTG 7149  
 Qy 5882 GAGTCACTCTCTGTTGGACTTCTCACCCAACTGGTATGCTCATGGGGTGGAGTTC 5941  
 Db 7150 GAGTCACTCTCTGTTGGACTTCTCACCCAACTGGTATGCTCATGGGGTGGAGTTC 7209  
 Qy 5942 AAGATCAGGCAAGAGAAACATGTAAGAAAGTAACTTCCCAACTCACCGGGTACATG 6001  
 Db 7210 AAGATCAGGCAAGAGAAACATGTAAGAAAGTAACTTCCCAACTCACCGGGTACATG 7269  
 Qy 6002 GCACCTTAGCCCTTCAAGAGATAGATCTCTCAAACTCATATGAAGCCTTCGTACCC 6061  
 Db 7270 GCACCTTAGCCCTTCAAGAGATAGATCTCTCAAACTCATATGAAGCCTTCGTACCC 7329  
 Qy 6062 ATACTCGCTGCTAGCCTATTATATACCACTCTCACTGGGCTCCATGAGTCTCGGCC 6121  
 Db 7330 ATACTCGCTGCTAGCCTATTATATACCACTCTCACTGGGCTCCATGAGTCTCGGCC 7389  
 Qy 6122 AAAACCTTAACTTGTGGATATGCTCCCTGAACTTTCARGCCATATGTTTCAATCC 6181  
 Db 7390 AAAACCTTAACTTGTGGATATGCTCCCTGAACTTTCARGCCATATGTTTCAATCC 7449  
 Qy 6182 CTGTACTTGAACAATGGAAACAATTTCAGACAGAAATAAACAACCTTCCCGTTTGTAG 6241  
 Db 7450 CTGTACTTGAACAATGGAAACAATTTCAGACAGAAATAAACAACCTTCCCGTTTGTAG 7509  
 Qy 6242 GACCTCTGTTTCCAAATGGAATAACCATACCTTCAACCTTCAACCTTGTGTAATTTA 6301  
 Db 7510 GACCTCTGTTTCCAAATGGAATAACCATACCTTCAACCTTCAACCTTGTGTAATTTA 7569  
 Qy 6302 GCAATACTACATACACAACAACCTTCCCAATGATCAGGTGGGTAACTCTCCCAACAAA 6361  
 Db 7570 GCAATACTACATACACAACAACCTTCCCAATGATCAGGTGGGTAACTCTCCCAACAAA 7629  
 Qy 6362 TAGTCTGCTTACCTCAGGAATAATTTTGTCTGTGGTACCTCAGCCTATCGTTTGTGA 6421  
 Db 7630 TAGTCTGCTTACCTCAGGAATAATTTTGTCTGTGGTACCTCAGCCTATCGTTTGTGA 7689  
 Qy 6422 ATGGCTCTTCAGAACTATGCTCTCTCTTCTAGTGGCCCTTATGRCATCTACA 6481  
 Db 7690 ATGGCTCTTCAGAACTATGCTCTCTCTTCTAGTGGCCCTTATGRCATCTACA 7749  
 Qy 6482 CTGAACAAGATTATACAGTTATGTCATATCTAAGCCCGGCAACAAAGAGTACCCATTC 6541  
 Db 7750 CTGAACAAGATTATACAGTTATGTCATATCTAAGCCCGGCAACAAAGAGTACCCATTC 7809

Qy 6542 TTCTTTTGTATTAGGACAGAGTGTAGTGCAGTACTAGTACTGGCATTTGGCGGTATCA 6601  
 Db 7810 TTCTTTTGTATTAGGACAGAGTGTAGTGCAGTACTAGTACTGGCATTTGGCGGTATCA 7869  
 Qy 6602 CAACCTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAATGGGACATGGAAACGG 6661  
 Db 7870 CAACCTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAATGGGACATGGAAACGG 7929  
 Qy 6662 TCGCGCACTCCCTGTCTACTTCAAGATCAACTTAACTTCCCTAGCAGCAGTAGTCTTTC 6721  
 Db 7930 TCGCGCACTCCCTGTCTACTTCAAGATCAACTTAACTTCCCTAGCAGCAGTAGTCTTTC 7989

RESULT 8

US-10-786-970A-21  
 ; Sequence 21, Application US/10786970A  
 ; Publication No. US2005006449A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JOAN, KNOLL  
 ; APPLICANT: ROGAN, PETER  
 ; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING  
 ; FILE REFERENCE: 30307  
 ; CURRENT APPLICATION NUMBER: US/10/786.970A  
 ; CURRENT FILING DATE: 2004-02-24  
 ; PRIOR APPLICATION NUMBER: US/09/573,080  
 ; PRIOR FILING DATE: 2000-05-16  
 ; NUMBER OF SEQ ID NOS: 479  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 21  
 ; LENGTH: 8523  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: repeat region  
 ; LOCATION: (1)..(8523)  
 ; OTHER INFORMATION: herv17  
 ; PUBLICATION INFORMATION:  
 ; PUBLICATION INFORMATION:  
 ; AUTHORS: Jurka, J, Walichiewicz, J, Milosavljevic, A  
 ; TITLE: Prototypic sequences for human repetitive DNA  
 ; JOURNAL: Journal of Molecular Evolution  
 ; VOLUME: 35  
 ; ISSUE: 4  
 ; PAGES: 286-291  
 ; DATE: 1992-10-  
 ; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)  
 ; DATABASE ENTRY DATE:  
 ; DATABASE ENTRY DATE: 1996-01-26  
 US-10-786-970A-21

Query Match 10.3%; Score 779; DB 10; Length 8523;  
 Best Local Similarity 98.9%; Pred. No. 0;  
 Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;  
 Qy 4802 AGGCTTCCCTGTAGGACAGAAAGGCCAGAGGTAATGAAGCACTAGTTCATGAA 4861  
 Db 6071 AGGCTTCCCTGTAGGACAGAAAGGCCAGAGGTAATGAAGCACTAGTTCATGAA 6130  
 Qy 4862 TAAATCCAGAGTTCCGCACTTCCCGAGGTTACAGAGTGACAATAGCCCTGCTTCCAGG 4921  
 Db 6131 TAAATCCAGAGTTCCGCACTTCCCGAGGTTACAGAGTGACAATAGCCCTGCTTCCAGG 6190  
 Qy 4922 CCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATAGTATACATTCATCTGCGCCT 4981  
 Db 6191 CCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATAGTATACATTCATCTGCGCCT 6250  
 Qy 4982 GAAGGCCACAGTCTCAGGAGGTCGAGAAATGAATGAAYACTCAAGAGCATCTAA 5041  
 Db 6251 GAAGGCCACAGTCTCAGGAGGTCGAGAAATGAATGAAYACTCAAGAGCATCTAA 6310  
 Qy 5042 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCCTGCTGTGCTATAGCCTTAAAAA 5101



Db 6311 AABAGCAAAACCCAGGAAACCACTCACAATGGCCCTGCTCTGTTGCCCTATAGCCTTAAAAA 6370  
 QY 5102 GAATCTGCAACTTTCCCAAAAAAGCAGGACTTACCCCATACGAATCTGTATGGAAGGC 5161  
 Db 6371 GAATCTGCAACTTTCCCAAAAAAGCAGGACTTACCCCATACGAATCTGTATGGAAGGC 6430  
 QY 5162 CTTTCATAACCAATGACCTTGTGCTTGCACCCAGAGCACTTATAGTTCGACAGCATCAC 5221  
 Db 6431 CTTTCATAACCAATGACCTTGTGCTTGCACCCAGAGCACTTATAGTTCGACAGCATCAC 6490  
 QY 5222 CTCCTTAGCCAAATATCAACAAGTTCTTAAAAACATTAACAAGAACCTTATCCCTGAGAGA 5281  
 Db 6491 CTCCTTAGCCAAATATCAACAAGTTCTTAAAAACATTAACAAGAACCTTATCCCTGAGAGA 6550  
 QY 5282 GGGAAAGAACTATTCCACCCWGTGACATGGTATTAAGTCAAGTCCCTTCYCTCAATTC 5341  
 Db 6551 GGGAAAGAACTATTCCACCCWGTGACATGGTATTAAGTCAAGTCCCTTCYCTCAATTC 6610  
 QY 5342 CCATCCCTAGATACATCTCGGGAAGGACCTACCCAGTCAATTTATATACCCCACTGC 5401  
 Db 6611 CCATCCCTAGATACATCTCGGGAAGGACCTACCCAGTCAATTTATATACCCCACTGC 6670  
 QY 5402 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACATTTAGTCAAAATCTTGGATACT 5461  
 Db 6671 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACATTTAGTCAAAATCTTGGATACT 6730  
 QY 5462 GCCAAAGAACCTGAAATCCAGAGACAAACGCTAGCTATTCTCTGTAACCTCTAGAGGA 5521  
 Db 6731 GCCAAAGAACCTGAAATCCAGAGACAAACGCTAGCTATTCTCTGTAACCTCTAGAGGA 6790  
 QY 5522 TTTGGCGCTGCTCTTCAACCAACAAACAGAGGAAGTAACTAAATCATAAATCCCCCA 5581  
 Db 6791 TTTGGCGCTGCTCTTCAACCAACAAACAGAGGAAGTAACTAAATCATAAATCCCCCA 6849  
 QY 5582 TGGSCCTCCCTTATCATATTTTCTCTKTAATGTTTAACTCTTCACTCTCACTG 5641  
 Db 6850 TGGSCCTCCCTTATCATATTTTCTCTKTAATGTTTAACTCTTCACTCTCACTG 6909  
 QY 5642 CACCCCTCCATGCGCTGATGACCACTAGTCTCCCTTACCMAGAGTTTCTATGAGAA 5701  
 Db 6910 CACCCCTCCATGCGCTGATGACCACTAGTCTCCCTTACCMAGAGTTTCTATGAGAA 6969  
 QY 5702 TGCAGGCTCCGGAATATGATGCGCCATCGATAGAGTCTTTTAAAGGGAACCCCA 5761  
 Db 6970 TGCAGGCTCCGGAATATGATGCGCCATCGATAGAGTCTTTTAAAGGGAACCCCA 7029  
 QY 5762 CTTCACTGCCACACCATATGCGCGGAACCTGCTATCACTCTGCACTCTTTGATGC 5821  
 Db 7030 CTTCACTGCCACACCATATGCGCGGAACCTGCTATCACTCTGCACTCTTTGATGC 7089  
 QY 5822 ATGCAAACTACTCATTTATGACAGGAAATGATTAATCTAGTCTGCTGGAGGACTTG 5881  
 Db 7090 ATGCAAACTACTCATTTATGACAGGAAATGATTAATCTAGTCTGCTGGAGGACTTG 7149  
 QY 5882 GAGTCACTGCTGTGGAATCTTCAACCAACTGATGCTGATGCTGATGCGGGTGGAGTTC 5941  
 Db 7150 GAGTCACTGCTGTGGAATCTTCAACCAACTGATGCTGATGCTGATGCGGGTGGAGTTC 7209  
 QY 5942 AAGATCAGCAAGAGAAACATGTAAGAAGTATCTCCCAACTCACCGGGGTACATG 6001  
 Db 7210 AAGATCAGCAAGAGAAACATGTAAGAAGTATCTCCCAACTCACCGGGGTACATG 7269  
 QY 6002 GCACCTTAGCCCTCAAAAGGACTAGTCTCTCAAAAGTATCTCAAAAGTATCTCCCAACTCACCGGGGTACATG 6061  
 Db 7270 GCACCTTAGCCCTCAAAAGGACTAGTCTCTCAAAAGTATCTCAAAAGTATCTCCCAACTCACCGGGGTACATG 7329  
 QY 6062 ATACTCGCTGGTAAAGCTATTTTAAATACCACTCTGAGTCTGAGTCTCGGCC 6121  
 Db 7330 ATACTCGCTGGTAAAGCTATTTTAAATACCACTCTGAGTCTGAGTCTCGGCC 7389  
 QY 6122 AAAACCTTACTAACTGTGATATGCTCCCTGAACTTTTCAAGCCATATGTTTCAATCC 6181  
 Db 7390 AAAACCTTACTAACTGTGATATGCTCCCTGAACTTTTCAAGCCATATGTTTCAATCC 7449

QY 6182 CTGTACTCTGAACAATGGAACAACTTACGACAGAAATAAACACCACTTCGTTTTAGTAG 6241  
 Db 7450 CTGTACTCTGAACAATGGAACAACTTACGACAGAAATAAACACCACTTCGTTTTAGTAG 7509  
 QY 6242 GACCTCTTGTTCCTAATSTGGAATAAACCCATACCTTCAAACTTCACTGTGTAAATTTA 6301  
 Db 7510 GACCTCTTGTTCCTAATSTGGAATAAACCCATACCTTCAAACTTCACTGTGTAAATTTA 7569  
 QY 6302 GCATACTACTACATACACAACTCCCAATGCAATCAGGTGGGTAACTTCTCCACACAAA 6361  
 Db 7570 GCATACTACTACATACACAACTCCCAATGCAATCAGGTGGGTAACTTCTCCACACAAA 7629  
 QY 6362 TAGTCTGCCCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCTATATCGTTGTTTGA 6421  
 Db 7630 TAGTCTGCCCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCTATATCGTTGTTTGA 7689  
 QY 6422 ATGGCTCTTCAGAACTATATGTGCTTCTCTCATCTTGTGCTCCCTATGACCATCTACA 6481  
 Db 7690 ATGGCTCTTCAGAACTATATGTGCTTCTCTCATCTTGTGCTCCCTATGACCATCTACA 7749  
 QY 6482 CTGAACAAGATTTATACAGTTATGTATATCTAAGCCCGCAACAAAGAGTACCCATTC 6541  
 Db 7750 CTGAACAAGATTTATACAGTTATGTATATCTAAGCCCGCAACAAAGAGTACCCATTC 7809  
 QY 6542 TTCTTTTGTATAGGAGCAGGAGTCTAGTGTGCTAGTGTGCTAGTGTGCGGTATCA 6601  
 Db 7810 TTCTTTTGTATAGGAGCAGGAGTCTAGTGTGCTAGTGTGCTAGTGTGCGGTATCA 7869  
 QY 6602 CAACCTCTACTAGTCTTACTACAACTATCTCAAGAACTAAATGGGAGCATGGAACGGG 6661  
 Db 7870 CAACCTCTACTAGTCTTACTACAACTATCTCAAGAACTAAATGGGAGCATGGAACGGG 7929  
 QY 6662 TGGCCGACCTCCCTGGTCACTTGCAGAGTCACTTAACCTCCCTAGCAGAGTGTCTTC 6721  
 Db 7930 TGGCCGACCTCCCTGGTCACTTGCAGAGTCACTTAACCTCCCTAGCAGAGTGTCTTC 7989

RESULT 9

US-11-028-539-3  
 ; Sequence 3, Application US/11028539  
 ; Publication No. US20050118573A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ALLIEL, Patrick  
 ; APPLICANT: PERIN, Jean-Pierre  
 ; APPLICANT: RIEGER, Francois  
 ; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HI  
 ; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS  
 ; FILE REFERENCE: 200936USOPCT  
 ; CURRENT APPLICATION NUMBER: US/11/028,539  
 ; PRIOR FILING DATE: 2005-01-05  
 ; PRIOR APPLICATION NUMBER: US/09/719,554  
 ; PRIOR FILING DATE: 2001-01-18  
 ; PRIOR APPLICATION NUMBER: PCT/FR99/01513  
 ; NUMBER OF SEQ ID NOS: 122  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 3  
 ; LENGTH: 10499  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-11-028-539-3

Query Match 10.3%; Score 779; DB 13; Length 10499;  
 Best Local Similarity 98.9%; Pred. No. 0;  
 Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;  
 QY 4802 AGGCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTATATAAGGACATAGTTCATGAA 4861  
 Db 7101 AGGCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTATATAAGGACATAGTTCATGAA 7160  
 QY 4862 TAATTCAGAGTTCGAGCTTCCCGAGGCTTACAGAGTCAATAGCCCTGCTTCCAGG 4921

Db 7161 TAAATCCAGATTGGAGTTCCCGAGGCTTACAGAGTGCACAAATAGCCCTTCCAGG 7220  
 Qy 4922 CCACAGTAAACCCAGGAGATCCCGAGGCTTGGGTATACGATATCACATTACATCGGCT 4981  
 Db 7221 CCACAGTAAACCCAGGAGATCCCGAGGCTTGGGTATACGATATCACATTACATCGGCT 7280  
 Qy 4982 GAAGGCCACAGTCCCTCAGGGAAGGTGAGAAAATGAATGAAAYACTCAAAAGGACATCTAA 5041  
 Db 7281 GAAGGCCACAGTCCCTCAGGGAAGGTGAGAAAATGAATGAAACACTCAAAAGGACATCTAA 7340  
 Qy 5042 AAAAGCAAAACCCAGGAAACCCACCTCACATGCGCTGCTGTGTCCTATAGCTTAAAAA 5101  
 Db 7341 AAAAGCAAAACCCAGGAAACCCACCTCACATGCGCTGCTGTGTCCTATAGCTTAAAAA 7400  
 Qy 5102 GAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGAAGGC 5161  
 Db 7401 GAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGAAGGC 7460  
 Qy 5162 CTTTCATTAACCAATGACCTTGTGCTTACCCAGGACAGCAACTTAGTTGAGAGATCAC 5221  
 Db 7461 CTTTCATTAACCAATGACCTTGTGCTTACCCAGGACAGCAACTTAGTTGAGAGATCAC 7520  
 Qy 5222 CTCCTTAGCCAAATATCAACAAGTTCTTAAACATTTACAAGGAACCTATCCCTGAGAAGA 5281  
 Db 7521 CTCCTTAGCCAAATATCAACAAGTTCTTAAACATTTACAAGGAACCTATCCCTGAGAAGA 7580  
 Qy 5282 GGGAAAGAACTATTCCACCWGTGACATGGTATTTAGTCAAGTCCCTTCYCTCTAATTC 5341  
 Db 7581 GGGAAAGAACTATTCCACCWGTGACATGGTATTTAGTCAAGTCCCTTCYCTCTAATTC 7640  
 Qy 5342 CCATCCCTTAGATACATCTGCGGAAGGACCCCTACCCAGTCAATTTTATATVACCCCAACTGC 5401  
 Db 7641 CCATCCCTTAGATACATCTGCGGAAGGACCCCTACCCAGTCAATTTTATATVACCCCAACTGC 7700  
 Qy 5402 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACATCTGAGTCAAACTCTGGATCT 5461  
 Db 7701 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACATCTGAGTCAAACTCTGGATCT 7760  
 Qy 5462 GCCAAAGGAACCTGAAATATCCAGGAGACAAACGCTAGCTATTCCTGTGAAACCTCTAGAGA 5521  
 Db 7761 GCCAAAGGAACCTGAAATATCCAGGAGACAAACGCTAGCTATTCCTGTGAAACCTCTAGAGA 7820  
 Qy 5522 TTGCGCCTGCTCTTCAACACACACCCAGGAGGAAGTAACTAAATCATATAATCCCCA 5581  
 Db 7821 TTGCGCCTGCTCTTCAACACACACCCAGGAGGAAGTAACTAAATCATATAAT - CCCC 7879  
 Qy 5582 TGGSCCTCCCTTATCATATTTTCTCTKTAAGTSTTTTACCTCTTCTACCTCTCACTG 5641  
 Db 7880 TGGSCCTCCCTTATCATATTTTCTCTTACTTCTTACCTCTTCTACCTCTCACTG 7939  
 Qy 5642 CACCCCTCCATGCGCTGTATGACAGTAGTCCCTTACCMAGAGTTTCTATGAGAA 5701  
 Db 7940 CACCCCTCCATGCGCTGTATGACAGTAGTCCCTTACCMAGAGTTTCTATGAGAA 7999  
 Qy 5702 TGCAGGCTCCGGAATATGATGCCCATGATAGGAGTCTTTTAAGGGAACCCCA 5761  
 Db 8000 TGCAGGCTCCGGAATATGATGCCCATGATAGGAGTCTTTTAAGGGAACCCCA 8059  
 Qy 5762 CTTTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTCTGCACTCTTTCATGC 5821  
 Db 8060 CTTTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTCTGCACTCTTTCATGC 8119  
 Qy 5822 ATGCAAAATCTATTATGGACAGGAAAATGAATTAATCCTAGTTGTCTCGGAGGACTTG 5881  
 Db 8120 ATGCAAAATCTATTATGGACAGGAAAATGAATTAATCCTAGTTGTCTCGGAGGACTTG 8179  
 Qy 5882 GAGTCACTGTGTGGACTTACTTCAACCAACTGGTATGCTGATGGGGTGGAGTTC 5941  
 Db 8180 GAGTCACTGTGTGGACTTACTTCAACCAACTGGTATGCTGATGGGGTGGAGTTC 8239  
 Qy 5942 AAGATCAGGCAAGAGAAAAACATGTAAGAGAGTAACTTCCCAACTCACCSSGGGTACATG 6001  
 Db 8240 AAGATCAGGCAAGAGAAAAACATGTAAGAGAGTAACTTCCCAACTCACCSSGGGTACATG 8299

Qy 6002 GCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCC 6061  
 Db 8300 GCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCC 8359  
 Qy 6062 ATACTCGCCTGGTAAAGCCCTATTTAATACCAACCTCAGTGGGCTCCATGAGGTCTCGGCC 6121  
 Db 8360 ATACTCGCCTGGTAAAGCCCTATTTAATACCAACCTCAGTGGGCTCCATGAGGTCTCGGCC 8419  
 Qy 6122 AAAACCTTACTAATCTGTTGGATATGCTCCCTCGAACTTCARGCCATATGTTTCAATCC 6181  
 Db 8420 AAAACCTTACTAATCTGTTGGATATGCTCCCTCGAACTTCARGCCATATGTTTCAATCC 8479  
 Qy 6182 CTGTACTGGAACATGGAACAACTTTCAGCAGAGAAATAAACACCACTTCCGTTTGTAGTAG 6241  
 Db 8480 CTGTACTGGAACATGGAACAACTTTCAGCAGAGAAATAAACACCACTTCCGTTTGTAGTAG 8539  
 Qy 6242 GACCTCTTGTTCCTCAATSTGGAATTAACCCATACCTCAAACTCACCCTGTGTAAATTTA 6301  
 Db 8540 GACCTCTTGTTCCTCAATSTGGAATTAACCCATACCTCAAACTCACCCTGTGTAAATTTA 8599  
 Qy 6302 GCAATFACTACATACACAACTCCCAATGCATCAGGTGGGTAACTCCTCCACACAAA 6361  
 Db 8600 GCAATFACTACATACACAACTCCCAATGCATCAGGTGGGTAACTCCTCCACACAAA 8659  
 Qy 6362 TAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCGTTTGTGA 6421  
 Db 8660 TAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCGTTTGTGA 8719  
 Qy 6422 ATGGCTCTTCAAGATCTATGTCTCTCATTTCTAGTGTCCCTCATAGTGTGATCCATCTACA 6481  
 Db 8720 ATGGCTCTTCAAGATCTATGTCTCTCATTTCTAGTGTCCCTCATAGTGTGATCCATCTACA 8779  
 Qy 6482 CTGAACAGATTTTATACAGTTATGTCAATCTAAGCCCCGCAACAAAAGAGTACCATTTC 6541  
 Db 8780 CTGAACAGATTTTATACAGTTATGTCAATCTAAGCCCCGCAACAAAAGAGTACCATTTC 8839  
 Qy 6542 TTCCTTTTGTATAGGAGCAGAGTGTAGTGTGACATAGGTACTGGCATTTGGCGGTATCA 6601  
 Db 8840 TTCCTTTTGTATAGGAGCAGAGTGTAGTGTGACATAGGTACTGGCATTTGGCGGTATCA 8899  
 Qy 6602 CAACCTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG 6661  
 Db 8900 CAACCTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG 8959  
 Qy 6662 TCGCGACTCCCTGGTCACTTCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTTC 6721  
 Db 8960 TCGCGACTCCCTGGTCACTTCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTTC 9019

RESULT 10

US-09-873-367C-81  
 ; Sequence 81, Application US/09873367C  
 ; Publication No. US20030165839A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Young, Paul  
 ; APPLICANT: Soppet, Daniel  
 ; APPLICANT: Endress, Gregory  
 ; APPLICANT: Augustus, Meena  
 ; APPLICANT: Ebner, Reinhard  
 ; APPLICANT: Carter, Kenneth  
 ; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using  
 ; FILE OF INVENTION: Signature Gene Sets  
 ; FILE REFERENCE: 689290-64  
 ; CURRENT APPLICATION NUMBER: US/09/873,367C  
 ; CURRENT FILING DATE: 2003-04-29  
 ; PRIOR APPLICATION NUMBER: U.S. 60/236,891  
 ; PRIOR FILING DATE: 2000-09-29  
 ; PRIOR APPLICATION NUMBER: U.S. 60/236,842  
 ; PRIOR FILING DATE: 2000-09-29  
 ; PRIOR APPLICATION NUMBER: U.S. 60/244,867  
 ; PRIOR FILING DATE: 2000-11-01  
 ; PRIOR APPLICATION NUMBER: U.S. 60/245,084

; PRIOR FILING DATE: 2000-11-01  
 ; NUMBER OF SEQ ID NOS: 1067  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 81  
 ; LENGTH: 56093  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-873-367C-81

Query Match		10.3%; Score 779; DB 3; Length 56093;
Best Local Similarity		98.9%; Pred. No. 0;
Matches 1899; Conservative		0; Mismatches 20; Indels 1; Gaps 1;

  

QY	4802	AGGCTTCCCTCTAGGACGAGAAAGCCCAAGAGTAAATAAGGCACTAGTTCATGAAA	4861
DB	35101	AGGCTTCCCTCTAGGACGAGAAAGCCCAAGAGTAAATAAGGCACTAGTTCATGAAA	35160
QY	4862	TAATTCACAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCAGG	4921
DB	35161	TAATTCACAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCAGG	35220
QY	4922	CCACAGTAACCCAGGGAGTATCCAGGGCTTACGATATCACTTACATCTGCGCCT	4981
DB	35221	CCACAGTAACCCAGGGAGTATCCAGGGCTTACGATATCACTTACATCTGCGCCT	35280
QY	4982	GAAGGCCACAGTCTCAGGGAAGGTGAGAAATGAATGAATACTCAAGGACATCTAA	5041
DB	35281	GAAGGCCACAGTCTCAGGGAAGGTGAGAAATGAATGAATACTCAAGGACATCTAA	35340
QY	5042	AAAAGCAAACCCAGGAAACCCACTCACATGGCTGTCTGTGCTTATAGCCCTTAAAAA	5101
DB	35341	AAAAGCAAACCCAGGAAACCCACTCACATGGCTGTCTGTGCTTATAGCCCTTAAAAA	35400
QY	5102	GAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAATGCTGTATGGAAGC	5161
DB	35401	GAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAATGCTGTATGGAAGC	35460
QY	5162	CCATTACCAATGACCTTGTGTTGACCCAGACAGCAACTTAGTTCGACATCAC	5221
DB	35461	CCATTACCAATGACCTTGTGTTGACCCAGACAGCAACTTAGTTCGACATCAC	35520
QY	5222	CTCCTTAGCCAAATATCAACAGTCTTAAACATTTACAGGAACCTATCCCTGAGAGA	5281
DB	35521	CTCCTTAGCCAAATATCAACAGTCTTAAACATTTACAGGAACCTATCCCTGAGAGA	35580
QY	5282	GGGAAAGAACTATTCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTC	5341
DB	35581	GGGAAAGAACTATTCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTC	35640
QY	5342	CCCATCCCTAGATACATCTCTGGGAAGACCCCTACCCAGTCAATTTATYTAACCCCACTGC	5401
DB	35641	CCCATCCCTAGATACATCTCTGGGAAGACCCCTACCCAGTCAATTTATYTAACCCCACTGC	35700
QY	5402	GGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTAGTCAATCTCGGATACT	5461
DB	35701	GGTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTAGTCAATCTCGGATACT	35760
QY	5462	GCCAAAGGAACCTGAAATCCAGAGACAACGCTAGCTATTCTGTGGAACCTCTAGAGGA	5521
DB	35761	GCCAAAGGAACCTGAAATCCAGAGACAACGCTAGCTATTCTGTGGAACCTCTAGAGGA	35820
QY	5522	TTTTCGCGCTGCTCTTCAAAACAACCCAGGAGAAAGTAACATAAAATCAATAATCCCCCA	5581
DB	35821	TTTTCGCGCTGCTCTTCAAAACAACCCAGGAGAAAGTAACATAAAATCAATAATCCCCCA	35879
QY	5582	TGSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTVACCTSTTTTCACTCTCACTG	5641
DB	35880	TGSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTVACCTSTTTTCACTCTCACTG	35939
QY	5642	CACCCCTCCCTGCTGTATGACACAGTAGCTCCCTVACCMAGAGTTTCTATGAGAA	5701
DB	35940	CACCCCTCCCTGCTGTATGACACAGTAGCTCCCTVACCMAGAGTTTCTATGAGAA	35999

RESULT 11

US-10-843-641A-81  
 ; Sequence 81, Application US/10843641A  
 ; Publication No. US200500644541

QY	5702	TGCAGCGTCCCGGAAATATTTGATGCCCATCTGTATAGGAGTCTTTSTAAGGGAACCCCA	5761
DB	36000	TGCAGCGTCCCGGAAATATTTGATGCCCATCTGTATAGGAGTCTTTSTAAGGGAACCCCA	36059
QY	5762	CCTTCACTGCCACACCCCATATGCCCCGCAACTGCTATCACTCTGCACTCTTTTGATGC	5821
DB	36060	CCTTCACTGCCACACCCCATATGCCCCGCAACTGCTATCACTCTGCACTCTTTTGATGC	36119
QY	5822	ATGCAAAATCTCATTTATGGACAGGAAATGATTAATCCTAGTTGCTCTGGAGGACTTG	5881
DB	36120	ATGCAAAATCTCATTTATGGACAGGAAATGATTAATCCTAGTTGCTCTGGAGGACTTG	36179
QY	5882	GAGTCACTGCTGTTGGACTTACTTCAACCAACTGCTATGCTATGCTGAGGGGTGGAGTTC	5941
DB	36180	GAGTCACTGCTGTTGGACTTACTTCAACCAACTGCTATGCTATGCTGAGGGGTGGAGTTC	36239
QY	5942	AGATCAGCAAGAGAGAAACATGTAAGAAAGTAATCTCCAACTCACCCSGGGTACATG	6001
DB	36240	AGATCAGCAAGAGAGAAACATGTAAGAAAGTAATCTCCAACTCACCCSGGGTACATG	36299
QY	6002	GCACTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCC	6061
DB	36300	GCACTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCC	36359
QY	6062	ATACTGCGCTGCTAAGCCCTATTTAATACCACTCTGCTGGCTCCATGAGGCTCTCGGCC	6121
DB	36360	ATACTGCGCTGCTAAGCCCTATTTAATACCACTCTGCTGGCTCCATGAGGCTCTCGGCC	36419
QY	6122	AAAAACCTACTAATCTGCTGATATGCTCCCTGAACTTTCARGCCATATGTTTCAATCC	6181
DB	36420	AAAAACCTACTAATCTGCTGATATGCTCCCTGAACTTTCARGCCATATGTTTCAATCC	36479
QY	6182	CTGTACTGAAACAATGGAACAACTTTCAGCACAGAAATAAACACCACTTCCGTTTATAGTAG	6241
DB	36480	CTGTACTGAAACAATGGAACAACTTTCAGCACAGAAATAAACACCACTTCCGTTTATAGTAG	36539
QY	6242	GACCTCTGTTTCCAAATSTGGAAATAACCCATACCTCAAACTCACTGCTGTAAAAATTTA	6301
DB	36540	GACCTCTGTTTCCAAATSTGGAAATAACCCATACCTCAAACTCACTGCTGTAAAAATTTA	36599
QY	6302	GCAATACATACATACACCAACTCCCAATGATCAGGTGGGTAACTCTCCACACAAA	6361
DB	36600	GCAATACATACATACACCAACTCCCAATGATCAGGTGGGTAACTCTCCACACAAA	36659
QY	6362	TAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTACCTCAGGCTATCGTTGTTTGA	6421
DB	36660	TAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTACCTCAGGCTATCGTTGTTTGA	36719
QY	6422	ATGGCTCTTTCAGAACTATGCTGCTCTCTCATTTCTAGTGCCCCCTATGRCCTATACA	6481
DB	36720	ATGGCTCTTTCAGAACTATGCTGCTCTCTCATTTCTAGTGCCCCCTATGRCCTATACA	36779
QY	6482	CTGAACAGATTTATACAGTTATGCTATGCTATGCTTAAAGCCCGCAACAAAGAGTACCATTC	6541
DB	36780	CTGAACAGATTTATACAGTTATGCTATGCTTAAAGCCCGCAACAAAGAGTACCATTC	36839
QY	6542	TTCTCTTTTCTATAGGACGAGGTGCTAGGTGCACTAGTACTGSCATTTGGCGGTATCA	6601
DB	36840	TTCTCTTTTCTATAGGACGAGGTGCTAGGTGCACTAGTACTGSCATTTGGCGGTATCA	36899
QY	6602	CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAATGGGACATGGAACGGG	6661
DB	36900	CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAATGGGACATGGAACGGG	36959
QY	6662	TGCGCGACTCCCTGGTCACTTGGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTTC	6721
DB	36960	TGCGCGACTCCCTGGTCACTTGGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTTC	37019

; GENERAL INFORMATION:

APPLICANT: Avalon Pharmaceuticals, Inc.

**TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using**

**TITLE OF INVENTION:** Signature Gene Sets

FILE REFERENCE: 689290-189

; CURRENT APPLICATION NUMBER: US/10

; CURRENT FILING DATE: 2004-05-12

; PRIOR APPLICATION NUMBER: US/09/873,367

; PRIOR FILING DATE: 2001-06-05

;; PRIOR APPLICATION NUMBER: US/09/9

; PRIOR FILING DATE: 2001-09-18

PRIOR APPLICATION NUMBER: US/09/0876

; PRIOR FILING DATE: 2001-09-25

PRIOR APPLICATION NUMBER: US/09/1

; PRIOR FILING DATE: 2001-09-25  
 ; PRIOR APPLICATION NUMBER: US/00/062 933

PRIOR APPLICATION NUMBER: US/09/

; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US/00/054 924

;; PRIOR APPLICATION NUMBER: US/09/77777777  
: PRIOR FILING DATE: 2001-09-27

; PRIOR FILING DATE: 2001-09-27  
 : PRIOR APPLICATION NUMBER: IIS/09/967 768

; PRIOR AFFILIATION NUMBER: US/057  
 : PRIOR FILING DATE: 2001-09-28

; PRIOR FILING DATE: 2001-09-28  
 ; PRIOR APPLICATION NUMBER: US/09/968.007

: PRIOR FILING DATE: 2001-10-02  
 : PRIOR APPLICATION NUMBER: 03/057

: PRIORITY DATE: 2001-10-02  
 : PRIORITY APPLICATION NUMBER: US/09/969,347

PRIOR FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: US/09/969,708

PRIOR FILING DATE: 2001-10-03

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 8447

```
; SOFTWARE: PatentIn version 3.0
```

; SEQ ID NO 81

; LENGTH: 56093

; TYPE: DNA

**ORGANISM: Homo sapiens**

US-10-843-641A-81

Query Match	10.3%;	Score 779;	DB 10;	Length 56093;
Best Local Similarity	98.9%;	Pred. No. 0;		
Matches 1899;	Conservative	0;	Mismatches 20;	Indels 1; Gaps 1;
QY	4802	AGGCCTTCCCTGTAGGACAGAAAAAGGCCCAAGAGGTAAATAAGGCATTAGTTTCATGAAA	4861	
DB	35101	AGGCCTTCCCTGTAGGACAGAAAAAGGCCCAAGAGGTAAATAAGGCATTAGTTTCATGAAA	35160	
QY	4862	TAAATCCAGATTCGACATTCCTCCGAGGCTTCACAGGTGACAAATAGCCCTGCTTTCGAGG	4921	
DB	35161	TAAATCCAGATTCGACATTCCTCCGAGGCTTCACAGGTGACAAATAGCCCTGCTTTCGAGG	35220	
QY	4922	CCACAGTAAACCCAGGGAGGTATCCGAGGCGTTAGGTATACGATATCACTTTACACTCGCGCT	4981	
DB	35221	CCACAGTAAACCCAGGGAGGTATCCGAGGCGTTAGGTATACGATATCACTTTACACTCGCGCT	35280	
QY	4982	GAAAGCCACAGTCCTCAGGGAAAGGTCGAGAAAATGAATGAAAYACTCAAAGGACATCTAA	5041	
DB	35281	GAAAGCCACAGTCCTCAGGGAAAGGTCGAGAAAATGAATGAAACACTCAAAGGACATCTAA	35340	
QY	5042	AAAAGCAAAACCCAGGAAAACCCACCTCACATGGCGCTGYCTGTTGGCTATAGCCTTAAAAA	5101	
DB	35341	AAAAGCAAAACCCAGGAAAACCCACCTCACATGGCGCTGYCTGTTGGCTATAGCCTTAAAAA	35400	
QY	5102	GAACTGTGCAACTTTTCCCAAAAAGCAGGAGCTTAGCCCATACGAAATGCTGTATGGAAGGC	5161	
DB	35401	GAACTGTGCAACTTTTCCCAAAAAGCAGGAGCTTAGCCCATACGAAATGCTGTATGGAAGGC	35460	
QY	5162	CCTTCATAAACCAATGACCTTGCTTGACCCCAAGACAGCCAACTTAGTTTGACAGCATCAC	5221	
DB	35461	CCTTCATAAACCAATGACCTTGCTTGACCCCAAGACAGCCAACTTAGTTTGACAGCATCAC	35520	
QY	5222	CTCCCTTAGCCCAATATCAACAGTCTTTAAAAACATTTACAGGAACTTATCCCTGAGAGA	5281	
DB	35521	CTCCCTTAGCCCAATATCAACAGTCTTTAAAAACATTTACAGGAACTTATCCCTGAGAGA	35580	
QY	5282	GGGAAAAGAACCTATTCCACCCMWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAAATTC	5341	



Query Match		9.6%;	Score 731;	DB 8;	Length 1617;	
Best Local Similarity		99.4%;	Pred. No. 0;			
Matches 1031;		Conservative	0;	Mismatches	6;	Indels 0; Gaps 0;
QY	5685	AGAGTTTCTATGGAGATGACGGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT	5744			
DB	105	AGAGTTTCTATGGAGATGACGGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT	164			
QY	5745	TTSTAAGGGACCCCACTTCACTGCCCCACACCCATATGCCCCCAACTGCTATCACTC	5804			
DB	165	TTCTAAGGGAACCCCACTTCACTGCCCCACACCCATATGCCCCCAACTGCTATCACTC	224			
QY	5805	TGCCACTTTTGCATGCAATCACTATTATTGGACAGAGAAATGATTAATCCTAG	5864			
DB	225	TGCCACTTTTGCATGCAATCACTATTATTGGACAGAGAAATGATTAATCCTAG	284			
QY	5865	TTGTCTCGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCAACCCAACTGGTATGTC	5924			
DB	285	TTGTCTCGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCAACCCAACTGGTATGTC	344			
QY	5925	TGATGGGGTGGAGTCAAGATCAGGCAAGAGAAAAACATGTAAAGAGTAATCTCCCA	5984			
DB	345	TGATGGGGTGGAGTCAAGATCAGGCAAGAGAAAAACATGTAAAGAGTAATCTCCCA	404			
QY	5985	ACTCACCSGGTACATGGCACCTCTAGCCCTACAAAGACTAGATCTCTCAAAACTACA	6044			
DB	405	ACTCACCSGGTACATGGCACCTCTAGCCCTACAAAGACTAGATCTCTCAAAACTACA	464			
QY	6045	TGAAACCCCTCCGTACCCATCTCGCTGGTAAGCCCTATTATACACCCCTCAGTGGGT	6104			
DB	465	TGAAACCCCTCCGTACCCATCTCGCTGGTAAGCCCTATTATACACCCCTCAGTGGGT	524			
QY	6105	CCATGAGTCTCGGCCAAACCCCTAATCTGTTGGATATGCCCTCCCGCTGAATTCAR	6164			
DB	525	CCATGAGTCTCGGCCAAACCCCTAATCTGTTGGATATGCCCTCCCGCTGAATTCAG	584			
QY	6165	GCATATGTTTCAATCCCTGTACCTGAACAATGGAAACAATTCAGACAGAAATAAACAC	6224			
DB	585	GCATATGTTTCAATCCCTGTACCTGAACAATGGAAACAATTCAGACAGAAATAAACAC	644			
QY	6225	CACCTTCGGTTTGTAGGACCTCTTGTGTTTCCAATSTGGAATAAACCCATACCTCAAACT	6284			
DB	645	CACCTTCGGTTTGTAGGACCTCTTGTGTTTCCAATCTGGAATAAACCCATACCTCAAACT	704			
QY	6285	CACCTGTGTAAATTTAGCAATACATACATACACCAACTCCCAATGCATCAGTGGGT	6344			
DB	705	CACCTGTGTAAATTTAGCAATACATACATACACCAACTCCCAATGCATCAGTGGGT	764			
QY	6345	AACTCTCCCAACAAATAGTCTGCTACCCCTCAGGAATATTTTGTGTTGGTACCTC	6404			
DB	765	AACTCTCCCAACAAATAGTCTGCTACCCCTCAGGAATATTTTGTGTTGGTACCTC	824			
QY	6405	AGCCTATCGTTGTTGAATGGCTCTTCAAGATCTATGCTTCTCTCATCTTAGTGCC	6464			
DB	825	AGCCTATCGTTGTTGAATGGCTCTTCAAGATCTATGCTTCTCTCATCTTAGTGCC	884			
QY	6465	CCCYATGRCATCTACATGAACAAGATTTATACAGTTATGATCATCTAAGCCCCGCAA	6524			
DB	885	CCCTATGACCATCTACATGAACAAGATTTATACAGTTATGATCATCTAAGCCCCGCAA	944			
QY	6525	CAAAAGAGTACCATCTCTCTTTTGTATAGGACGAGGTGCTAGGTGCACTAGGTAC	6584			
DB	945	CAAAAGAGTACCATCTCTCTTTTGTATAGGACGAGGTGCTAGGTGCACTAGGTAC	1004			
QY	6585	TGGCAATGGCGGTATCACAACTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAA	6644			
DB	1005	TGGCAATGGCGGTATCACAACTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAA	1064			
QY	6645	TGGGACATGGAACGGGTGCGGACTCCCTGGTCACTTGGCAAGATCAACTTAATCCCT	6704			
DB	1065	TGGGACATGGAACGGGTGCGGACTCCCTGGTCACTTGGCAAGATCAACTTAATCCCT	1124			
QY	6705	AGCAGCAGTACCTTTC 6721				

DB	1125	AGCAGCAGTACTCTTC	1141			
RESULT 14						
US-10-632-793-24						
; Sequence 24, Application US/10632793						
; Publication No. US20040048298A1						
; GENERAL INFORMATION:						
; APPLICANT: PARANHOS-BACCALA, Glaucia						
; APPLICANT: MALLET, Francois						
; APPLICANT: VOISSET, Cecile						
; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN						
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT						
; FILE REFERENCE: 110048						
; CURRENT APPLICATION NUMBER: US/10/632,793						
; CURRENT FILING DATE: 2003-08-04						
; PRIOR APPLICATION NUMBER: US/09/869,927						
; PRIOR FILING DATE: 2001-10-22						
; PRIOR APPLICATION NUMBER: PCT/FR00/00144						
; PRIOR FILING DATE: 2000-01-21						
; PRIOR APPLICATION NUMBER: FR 99/00888						
; PRIOR FILING DATE: 1999-01-21						
; NUMBER OF SEQ ID NOS: 33						
; SOFTWARE: PatentIn version 3.1						
; SEQ ID NO 24						
; LENGTH: 1948						
; TYPE: DNA						
; ORGANISM: Homo sapiens						
; FEATURE:						
; NAME/KEY: misc feature						
; LOCATION: (84)..(84)						
; OTHER INFORMATION: n = a or g or c or t/u						
; FEATURE:						
; NAME/KEY: misc feature						
; LOCATION: (193)..(193)						
; OTHER INFORMATION: n = a or g or c or t/u						
; FEATURE:						
; NAME/KEY: misc feature						
; LOCATION: (241)..(241)						
; OTHER INFORMATION: n = a or g or c or t/u						
US-10-632-793-24						
Query Match						
Best Local Similarity 9.6%; Score 731; DB 8; Length 1948;						
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;						
QY	5685	AGAGTTTCTATGGAGATGACGGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT	5744			
DB	823	AGAGTTTCTATGGAGATGACGGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT	882			
QY	5745	TTSTAAGGGAACCCCACTTCACTGCCCCACACCCATATGCCCCCAACTGCTATCACTC	5804			
DB	883	TTGTAAGGGAACCCCACTTCACTGCCCCACACCCATATGCCCCCAACTGCTATCACTC	942			
QY	5805	TGCCACTTTTGCATGCAATCACTATTATTGGACAGAGAAATGATTAATCCTAG	5864			
DB	943	TGCCACTTTTGCATGCAATCACTATTATTGGACAGAGAAATGATTAATCCTAG	1002			
QY	5865	TTGTCTCGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCAACCCAACTGGTATGTC	5924			
DB	1003	TTGTCTCGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCAACCCAACTGGTATGTC	1062			
QY	5925	TGATGGGGTGGAGTCAAGATCAGGCAAGAGAAAAACATGTAAAGAGTAATCTCCCA	5984			
DB	1063	TGATGGGGTGGAGTCAAGATCAGGCAAGAGAAAAACATGTAAAGAGTAATCTCCCA	1122			
QY	5985	ACTCACCSGGTACATGGCACTCTAGCCCTACAAAGACTAGATCTCTCAAAACTACA	6044			
DB	1123	ACTCACCSGGTACATGGCACTCTAGCCCTACAAAGACTAGATCTCTCAAAACTACA	1182			
QY	6045	TGAAACCCCTCCGTACCCATCTCGCTGGTAAGCCCTATTATACACCCCTCAGTGGGT	6104			



```

1183 TGAACCCCTCGTACCATCTACTCGCTGGTAAGCTATTTAATACACCCCTCACTGGCT 1242
6105 CCATGAGGTCTCGGCCCAAAACCTCTACTAAGTGTGGATATGCTCCCTGAACTTCAR 6164
1243 CCATGAGGTCTCGGCCCAAAACCTCTACTAAGTGTGGATATGCTCCCTGAACTTCAG 1302
6165 GCCATATGTTTCAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6224
1303 GCCATATGTTTCAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1362
6225 CACTTCGGTTTATGAGGACCTCTGTTTTCCTTCAATSTGGAATAAACCATACCTCAAACCT 6284
1363 CACTTCGGTTTATGAGGACCTCTGTTTTCCTTCAATSTGGAATAAACCATACCTCAAACCT 1422
6285 CACTTCGGTTTATGAGGACCTCTGTTTTCCTTCAATSTGGAATAAACCATACCTCAAACCT 6344
1423 CACTTCGGTTTATGAGGACCTCTGTTTTCCTTCAATSTGGAATAAACCATACCTCAAACCT 1482
6345 AACTCTCCCAACAAATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6404
1483 AACTCTCCCAACAAATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1542
6405 AGCTATCGTGTGTTGAATGGCTCTTCAAGATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 6464
1543 AGCTATCGTGTGTTGAATGGCTCTTCAAGATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1602
6465 CCCATGRCATCTACGAAACAGATTTATACAGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6524
1603 CCTATGCGCATCTACATGAAACAGATTTATACAGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1662
6525 CAAAGAGTACCAATCTCTCTTTTGTATAGGAGGAGTGTAGTGTGCTGCTGCTGCTGCTGCTGCT 6584
1663 CAAAGAGTACCAATCTCTCTTTTGTATAGGAGGAGTGTAGTGTGCTGCTGCTGCTGCTGCTGCT 1722
6585 TGCCATTTGGCGGTATCACACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA 6644
1723 TGCCATTTGGCGGTATCACACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA 1782
6645 TGCGGACATGGAACGGGTGCGGACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6704
1783 TGCGGACATGGAACGGGTGCGGACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1842
6705 AGCAGCAGTGTCTTCT 6721
1843 AGCAGCAGTGTCTTCT 1859

RESULT 15
US-10-717-580-5
; Sequence 5, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
; APPLICANT: BESEME, Frederic
; APPLICANT: BLOND, Jean-Luc
; APPLICANT: BOUTON, Olivier
; APPLICANT: MANDRAND, Bernard
; APPLICANT: MALLE, Hervé
; APPLICANT: PERRON, Hervé
; TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISEASES
; FILE REFERENCE: 105045
; CURRENT APPLICATION NUMBER: US/10717580
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US/09/446,024A
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/FR98/01442
; PRIOR FILING DATE: 1998-07-06
; PRIOR APPLICATION NUMBER: FR 97/08815
; PRIOR FILING DATE: 1997-07-07
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1948

```

```

; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84)..(84)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (193)..(193)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (241)..(241)
; OTHER INFORMATION: n = any nucleotide
;
US-10-717-580-5

Query Match      9.6%; Score 731; DB 9; Length 1948;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGAGAAATGAGCGTCCCGGAAATATTTGATGCCCCATCGTATAGGAGTCT 5744
DB 823 AGAGTTTCTATGAGAAATGAGCGTCCCGGAAATATTTGATGCCCCATCGTATAGGAGTCT 882
QY 5745 TTSTANGGAAACCCCAACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 5804
DB 883 TTGTAAAGGAAACCCCAACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 942
QY 5805 TGCCACTCTTTGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 5864
DB 943 TGCCACTCTTTGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1002
QY 5865 TTGTCTCGGAGGACTTTGGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5924
DB 1003 TTGTCTCGGAGGACTTTGGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1062
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5984
DB 1063 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1122
QY 5985 ACTCACCGGGTACATGCGACCTCTAGCCCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6044
DB 1123 ACTCACCGGGTACATGCGACCTCTAGCCCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1182
QY 6045 TGAACCCCTCGTACCCATCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6104
DB 1183 TGAACCCCTCGTACCCATCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1242
QY 6105 CCATGAGGTCTCGGCCCAAAACCTCTACTAAGTGTGGATATGCTCCCTGAACTTCAR 6164
DB 1243 CCATGAGGTCTCGGCCCAAAACCTCTACTAAGTGTGGATATGCTCCCTGAACTTCAG 1302
QY 6165 GCATATGTTTCAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6224
DB 1303 GCATATGTTTCAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1362
QY 6225 CACTTCGGTTTATGAGGACCTCTGTTTTCCTTCAATSTGGAATAAACCATACCTCAAACCT 6284
DB 1363 CACTTCGGTTTATGAGGACCTCTGTTTTCCTTCAATSTGGAATAAACCATACCTCAAACCT 1422
QY 6285 CACTTCGGTTTATGAGGACCTCTGTTTTCCTTCAATSTGGAATAAACCATACCTCAAACCT 6344
DB 1423 CACTTCGGTTTATGAGGACCTCTGTTTTCCTTCAATSTGGAATAAACCATACCTCAAACCT 1482
QY 6345 AACTCTCCCAACAAATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6404
DB 1483 AACTCTCCCAACAAATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1542
QY 6405 AGCTATCGTGTGTTGAATGGCTCTTCAAGATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6464
DB 1543 AGCTATCGTGTGTTGAATGGCTCTTCAAGATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1602
QY 6465 CCCATGRCATCTACGAAACAGATTTATACAGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6524

```



```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 2782
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-632-793-26

Query Match          9.6%; Score 731; DB 8; Length 2782;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGAGAAATGACGCGTCCCGGAAATATGATGCCCATCGTATAGAGTCT 5744
DB 867 AGAGTTTCTATGAGAAATGACGCGTCCCGGAAATATGATGCCCATCGTATAGAGTCT 926
QY 5745 TTSTAAGGGAAACCCCACTTCACTGCGCCACCAATATGCCCGCAACTGCTATCACTC 5804
DB 927 TTCTAAGGGAAACCCCACTTCACTGCGCCACCAATATGCCCGCAACTGCTATCACTC 986
QY 5805 TGCCACTCTTTGCATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5864
DB 987 TGCCACTCTTTGCATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1046
QY 5865 TTGTCCTCGAGGACTTGGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5924
DB 1047 TTGTCCTCGAGGACTTGGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAGAAACATGTAAAGAAATGTAATCTCCCA 5984
DB 1107 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAGAAACATGTAAAGAAATGTAATCTCCCA 1166
QY 5985 ACTCACCGGGGTACATGCGACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACA 6044
DB 1167 ACTCACCGGGGTACATGCGACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACA 1226
QY 6045 TGAAACCCCTCGTACCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6104
DB 1227 TGAAACCCCTCGTACCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1286
QY 6105 CCATGAGTCTCGGCGCAAAACCCCTACTAACTGTGTGATATGCTTCCCTCGAATCTTCA 6164
DB 1287 CCATGAGTCTCGGCGCAAAACCCCTACTAACTGTGTGATATGCTTCCCTCGAATCTTCA 1346
QY 6165 GCATATGTTTCAATCTCTGTACCTGAACTGAACTGCAAGGACTAGATCTCTCAAAACTACA 6224
DB 1347 GCATATGTTTCAATCTCTGTACCTGAACTGAACTGCAAGGACTAGATCTCTCAAAACTACA 1406
QY 6225 CACTCCGTTTGTAGTAGGACCTCTGTTTCAATGTTGGATATGCTTCCCTCGAATCTTCA 6284
DB 1407 CACTCCGTTTGTAGTAGGACCTCTGTTTCAATGTTGGATATGCTTCCCTCGAATCTTCA 1466
QY 6285 CACTCTGTAAATTTAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6344
DB 1467 CACTCTGTAAATTTAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1526
QY 6345 AACTCTCCCAACAAATAGTCTGCTACCTCCAGGAAATATTTTGTGCTGTGATGCTGCT 6404
DB 1527 AACTCTCCCAACAAATAGTCTGCTACCTCCAGGAAATATTTTGTGCTGTGATGCTGCT 1586
QY 6405 AGCCTATCGTTGTTGAATGGCTCTTCAAGATCTATGCTTCTCTCATTTCTAGTGCC 6464
DB 1587 AGCCTATCGTTGTTGAATGGCTCTTCAAGATCTATGCTTCTCTCATTTCTAGTGCC 1646
QY 6465 CCCTATGRCATCTACACTGAAAGATTTTATACAGTTTATGCTATATCTAAGCCCGCCAA 6524
DB 1647 CCCTATGRCATCTACACTGAAAGATTTTATACAGTTTATGCTATATCTAAGCCCGCCAA 1706
QY 6525 CAAAGAGTACCCATCTTCTCTTTTGTATAGGAGCAGGAGTGTAGTGCTACCTAGTGATC 6584
DB 1707 CAAAGAGTACCCATCTTCTCTTTTGTATAGGAGCAGGAGTGTAGTGCTACCTAGTGATC 1766
QY 6585 TGCGCATTTGGCGGTATCACAACTCTACTGATTTCTACTACAACTATCTCAAGAACTAAA 6644

```

```

1407 CACTTCCGTTTGTAGTAGACCTCTTGTTCGAATGTGGAATTAACCCATACCTCAACCT 1466
6285 CACTGTGTAAATTTAGCAATACATACATACACAAACCACTCCCAATGCATCAGGTGGGT 6344
1467 CACCTGTGTAAATTTAGCAATACATACATACACAAACCACTCCCAATGCATCAGGTGGGT 1526
6345 AACTCTCCACACAAATAGTCTGCTACCTCAGCAATATTTTGTCTGTGTACCTC 6404
1527 AACTCTCCACACAAATAGTCTGCTACCTCAGCAATATTTTGTCTGTGTACCTC 1586
6405 AGCTATCGTTTGTAAATGGCTCTTCAAGATCTAGTCTCTCATCTTCTAGTGCC 6464
1587 AGCTATCGTTTGTAAATGGCTCTTCAAGATCTAGTCTCTCATCTTCTAGTGCC 1646
6465 CCCATGRCATCTACCTGAAACAAAGATTTATACAGTTATGTCAATCTAAGCCCGCAA 6524
1647 CCCTATGACCATCTACCTGAAACAAAGATTTATACAGTTATGTCAATCTAAGCCCGCAA 1706
6525 CAAAAGAGTACCATCTCTCTTTTGTATAGGACGAGGTGCTAGGTGCATAGGTAC 6584
1707 CAAAAGAGTACCATCTCTCTTTTGTATAGGACGAGGTGCTAGGTGCATAGGTAC 1766
6585 TGGCAATGGCGGTATCAGAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 6644
1767 TGGCAATGGCGGTATCAGAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 1826
6645 TGGGACATGGAACGGGTGCGGACTCCCTGTGTCACCTTGGCAAGATCAACTTTAATCCCT 6704
1827 TGGGACATGGAACGGGTGCGGACTCCCTGTGTCACCTTGGCAAGATCAACTTTAATCCCT 1886
6705 AGCAGAGTAGTCTTTC 6721
1887 AGCAGAGTAGTCTTTC 1903

```

RESULT 19

```

US-10-717-580-7
; Sequence 7, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
; APPLICANT: BESEME, Frederic
; APPLICANT: BLOND, Jean-Luc
; APPLICANT: BOUTON, Olivier
; APPLICANT: MANDRAND, Bernard
; APPLICANT: MALLET, Francois
; APPLICANT: PERRON, Hervé
; TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISEASES
; TITLE OF INVENTION: WITH PREGNANCY DISORDERS
; FILE REFERENCE: 105045
; CURRENT APPLICATION NUMBER: US/10/717,580
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US/09/446,024A
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/FR98/01442
; PRIOR FILING DATE: 1998-07-06
; PRIOR APPLICATION NUMBER: FR 97/08815
; PRIOR FILING DATE: 1997-07-07
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 2782
; TYPE: DNA
; ORGANISM: Human
US-10-717-580-7

```

```

Query Match 9.6%; Score 731; DB 9; Length 2782;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 5685 AGAGTTTCTATGAGAAATCAGCGTCCCGAAATATTTGATGCCCCCATCGTATAGGAGTCT 5744
DB 867 AGAGTTTCTATGAGAAATCAGCGTCCCGAAATATTTGATGCCCCCATCGTATAGGAGTCT 926

```

```

QY 5745 TTSTAAGGGAACCCACCTTCACTGCCACACCCATATGCCGCACTGCTATCACTC 5804
DB 927 TTCTAAGGGAACCCACCTTCACTGCCACACCCATATGCCGCACTGCTATCACTC 986
QY 5805 TGCCACCTCTTGGCATGCAATATCTATTTATGGACAGGAAATATGATTAATCTCTAG 5864
DB 987 TGCCACCTCTTGGCATGCAATATCTATTTATGGACAGGAAATATGATTAATCTCTAG 1046
QY 5865 TTGCTCTGGAGGACTTGGAGTCACTGTCTGTGTGGACTTTACCCCAAACTGGTATGTC 5924
DB 1047 TTGCTCTGGAGGACTTGGAGTCACTGTCTGTGTGGACTTTACCCCAAACTGGTATGTC 1106
QY 5925 TGATGGGGGTGGAGTTCAGATCAGGCAAGAGAAACATGTTAAAGAGTAATCTCCCA 5984
DB 1107 TGATGGGGGTGGAGTTCAGATCAGGCAAGAGAAACATGTTAAAGAGTAATCTCCCA 1166
QY 5985 ACTCACGSGGTGATATGGCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACA 6044
DB 1167 ACTCACGSGGTGATATGGCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACA 1226
QY 6045 TGAACCCCTCCGTACCCATATCTCGCTGGTAAGCCTATTTAATACCACTCATCTGGGT 6104
DB 1227 TGAACCCCTCCGTACCCATATCTCGCTGGTAAGCCTATTTAATACCACTCATCTGGGT 1286
QY 6105 CCATGAGTCTCGGCGCCAAACCCCTACTAACTGTGGATATGCCCTCCCTGAACTTCAR 6164
DB 1287 CCATGAGTCTCGGCGCCAAACCCCTACTAACTGTGGATATGCCCTCCCTGAACTTCAR 1346
QY 6165 GCCATATGTTTCAATCCCTGTACCTGAAACAATGGAACAACCTTCAGCACAGAAATAAACAC 6224
DB 1347 GCCATATGTTTCAATCCCTGTACCTGAAACAATGGAACAACCTTCAGCACAGAAATAAACAC 1406
QY 6225 CACTTCCGTTTGTAGTAGGACCTCTTGTTCCAATSTGGAATAAACCCATACCTCAAACT 6284
DB 1407 CACTTCCGTTTGTAGTAGGACCTCTTGTTCCAATSTGGAATAAACCCATACCTCAAACT 1466
QY 6285 CACTTGTGTAATTTTAGCAATATCTACATACAAACCACTCCCAATGCATCAGGTGGGT 6344
DB 1467 CACTTGTGTAATTTTAGCAATATCTACATACAAACCACTCCCAATGCATCAGGTGGGT 1526
QY 6345 AACTCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTC 6404
DB 1527 AACTCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTC 1586
QY 6405 AGCTATCGTTTGTAAATGGCTCTTCAAGATCTATGTGCTCTCTCATCTTCTAGTGCC 6464
DB 1587 AGCTATCGTTTGTAAATGGCTCTTCAAGATCTATGTGCTCTCTCATCTTCTAGTGCC 1646
QY 6465 CCCTATGRCATCTACCTGAAACAAGATTTATACAGTTATGTCAATCTAAGCCCGCAA 6524
DB 1647 CCCTATGRCATCTACCTGAAACAAGATTTATACAGTTATGTCAATCTAAGCCCGCAA 1706
QY 6525 CAAAAGAGTACCCATCTTCTTTTGTATAGGACGAGGTGCTAGGTGCATAGGTAC 6584
DB 1707 CAAAAGAGTACCCATCTTCTTTTGTATAGGACGAGGTGCTAGGTGCATAGGTAC 1766
QY 6585 TGGCAATGGCGGTATCAGAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 6644
DB 1767 TGGCAATGGCGGTATCAGAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 1826
QY 6645 TGGGACATGGAACGGGTGCGGACTCCCTGTGTCACCTTGGCAAGATCAACTTTAATCCCT 6704
DB 1827 TGGGACATGGAACGGGTGCGGACTCCCTGTGTCACCTTGGCAAGATCAACTTTAATCCCT 1886
QY 6705 AGCAGAGTAGTCTTTC 6721
DB 1887 AGCAGAGTAGTCTTTC 1903

```

```

RESULT 20
US-09-902-535-1
; Sequence 1, Application US/09902535

```

```
; Patent No. US20020102530A1
; GENERAL INFORMATION:
; APPLICANT: Keith, Jr., James C.
; APPLICANT: McCoy, John M.
; APPLICANT: MI, Sha
; TITLE OF INVENTION: Methods and compositions for diagnosing
; TITLE OF INVENTION: and treating preclampsia and gestational trophoblast
; TITLE OF INVENTION: disorders
; FILE REFERENCE: GIN-6008B4
; CURRENT APPLICATION NUMBER: US/09/902,535
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,657
; PRIOR FILING DATE: 2000-07-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2930
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (930)...(2546)
; US-09-902-535-1

Query Match          9.6%; Score 728; DB 3; Length 2930;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1898; Conservative 0; Mismatches 21; Indels 1; Gaps 1;

QY 4802 AGGCTTCCCTCTAGGACAGAAAAGGCCCAAGAGGTAAATAAAGGCACTAGTTTCATGAAA 4861
DB 152 AGGCTTCCCTCTAGGACAGAAAAGGCCCAAGAGGTAAATAAAGGCACTAGTTTCATGAAA 211
QY 4862 TAATTCAGATTCGACTTCCCGAGGCTTACAGAGTGAATAGCCCTGCTTTCCAGG 4921
DB 212 TAATTCAGATTCGACTTCCCGAGGCTTACAGAGTGAATAGCCCTGCTTTCCAGG 271
QY 4922 CCACAGTAACCCAGGAGATGCCAGGCTTGGATATACATATCATCTTACATCGGCT 4981
DB 272 CCACAGTAACCCAGGAGATGCCAGGCTTGGATATACATATCATCTTACATCGGCT 331
QY 4982 GAAGGCCACAGTCTCAGGAGAGGTGAGAAAATGAATGAATACTCAAGGACATCTAA 5041
DB 332 GAAGGCCACAGTCTCAGGAGAGGTGAGAAAATGAATGAATACTCAAGGACATCTAA 391
QY 5042 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCTGTCTGTGGCTATAGCCCTTAAAAA 5101
DB 392 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCTGTCTGTGGCTATAGCCCTTAAAAA 451
QY 5102 GAATCTGCAACTTCCCAAAAAGGAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 5161
DB 452 GAATCTGCAACTTCCCAAAAAGGAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 511
QY 5162 CCTTCAATACCAATGACTTGTGCTTGACCAAGACAGCCCACTTAGTTGCGACATCAC 5221
DB 512 CCTTCAATACCAATGACTTGTGCTTGACCAAGACAGCCCACTTAGTTGCGACATCAC 571
QY 5222 CTCCTTAGCCAAATATCAACAAGTCTTAAAAATTTACAGGAAACCTATCCCTTGAGAAGA 5281
DB 572 CTCCTTAGCCAAATATCAACAAGTCTTAAAAATTTACAGGAAACCTATCCCTTGAGAAGA 631
QY 5282 GGGAAAAGAACTATTCACCCWGTGACATGGTATTTAGTCAAGTCCCTTCYCTCTAATTC 5341
DB 632 GGGAAAAGAACTATTCACCCWGTGACATGGTATTTAGTCAAGTCCCTTCYCTCTAATTC 691
QY 5342 CCCATCCCTAGATACATCTTGGGAAGGACCTTACCAGTCAATTTATATACCCCACTGC 5401
DB 692 CCCATCCCTAGATACATCTTGGGAAGGACCTTACCAGTCAATTTATATACCCCACTGC 751
QY 5402 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACCTTGGATCAAAATCCTGGATCT 5461
DB 752 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACCTTGGATCAAAATCCTGGATCT 811
QY 5462 GCCAAAGGAACTGAAAAATCCAGGAGACAAACGCTAGCTATTCTGTGAAACCTCTAGAGGA 5521
```

```
DB 812 GCCAAAGGAACTGAAAAATCCAGGAGACAAACGCTAGCTATTCTGTGAAACCTCTAGAGGA 871
QY 5522 TTTGCGCCTGCTCTTCAAAACAAACACAGGAGGAAAGTAACTAAATATCAATAATCCCCCA 5581
DB 872 TTTGCGCCTGCTCTTCAAAACAAACACAGGAGGAAAGTAACTAAATATCAATAATCCCCCA 930
QY 5582 TGGSCCTCCCTTATCATATTTTCTCTKTASTGTTTSTTTTACCTCTTTCACCTCTCATCTG 5641
DB 931 TGGCCCTCCCTTATCATATTTTCTCTTACTTCTTTTACCTCTTTCACCTCTCATCTG 990
QY 5642 CACCCCTCCATGCGCCTGTATGACACAGTAGCTCCCTTACCAAGAGTCTTCTATGAGAA 5701
DB 991 CACCCCTCCATGCGCCTGTATGACACAGTAGCTCCCTTACCAAGAGTCTTCTATGAGAA 1050
QY 5702 TGCAGCGTCCCGGAAATATTTGATGCCCATCGTATAGGAGTCTTCTTAAGGGAACCCCCA 5761
DB 1051 TGCAGCGTCCCGGAAATATTTGATGCCCATCGTATAGGAGTCTTCTTAAGGGAACCCCCA 1110
QY 5762 CTTTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTTGCATGC 5821
DB 1111 CTTTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTTGCATGC 1170
QY 5822 ATGCAATACTCATTTATGACAGGAAATATGATTAATCCTAGTTCTCTGAGGAGCTTG 5881
DB 1171 ATGCAATACTCATTTATGACAGGAAATATGATTAATCCTAGTTCTCTGAGGAGCTTG 1230
QY 5882 GAGTCACTGCTGTTGGACTTACTTCAACCAACTGCTATGCTGATGCGGGTGGAGTTC 5941
DB 1231 GAGTCACTGCTGTTGGACTTACTTCAACCAACTGCTATGCTGATGCGGGTGGAGTTC 1290
QY 5942 AAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCAACTCACCGGGTACATG 6001
DB 1291 AAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCAACTCACCGGGTACATG 1350
QY 6002 GCACCTCTAGCCCTTCAAAAGGACTAGATCTCTCAAACTCATATGAAACCTCTCGTACCC 6061
DB 1351 GCACCTCTAGCCCTTCAAAAGGACTAGATCTCTCAAACTCATATGAAACCTCTCGTACCC 1410
QY 6062 ATACTCGCTGTTAGGCTTATTAATACCACTCTGCGGCTCAATGAGTCTCGGCT 6121
DB 1411 ATACTCGCTGTTAGGCTTATTAATACCACTCTGCGGCTCAATGAGTCTCGGCT 1470
QY 6122 AAAACCTCTACTGTTGGATATGCTCCCTGAACTTTCARGCCATATGTTTCAATCC 6181
DB 1471 AAAACCTCTACTGTTGGATATGCTCCCTGAACTTTCARGCCATATGTTTCAATCC 1530
QY 6182 CTGTACTGAAACATGGAACAACTTCAGCACAGAAATAAACACCACTTCGGTTTATGATG 6241
DB 1531 CTGTACTGAAACATGGAACAACTTCAGCACAGAAATAAACACCACTTCGGTTTATGATG 1590
QY 6242 GACCTCTGTTTCCCAATSTGGAAATAAACCCATACCTCAACCTCACCTGTTGTAATAATTA 6301
DB 1591 GACCTCTGTTTCCCAATSTGGAAATAAACCCATACCTCAACCTCACCTGTTGTAATAATTA 1650
QY 6302 GCAATACTACATACACAACTCCCAATGCAATGAGTGGGTAACTCTCCCAACAA 6361
DB 1651 GCAATACTACATACACAACTCCCAATGCAATGAGTGGGTAACTCTCCCAACAA 1710
QY 6362 TAGTCTGCCCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGGCTATCGTTGTTGA 6421
DB 1711 TAGTCTGCCCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGGCTATCGTTGTTGA 1770
QY 6422 ATGGCTCTTCAAACTATGCTCTCTCTCATCTTAGTGGCCCTCATGTCATCTACA 6481
DB 1771 ATGGCTCTTCAAACTATGCTCTCTCTCATCTTAGTGGCCCTCATGTCATCTACA 1830
QY 6482 CTGAAACAGATTTATACAGTTATGATCATATCTAAGCCCGCAACAAAGAGTACCCATTC 6541
DB 1831 CTGAAACAGATTTATACAGTTATGATCATATCTAAGCCCGCAACAAAGAGTACCCATTC 1890
QY 6542 TTCTCTTTTGTATAGGAGCAGGAGTGTAGGTGCACTAGGTACTGTGCGGCTATCA 6601
```

Db 1891 TTCTTTTGTATTAGAGCAGGAGTGCTAGGTGCACCTAGGTACTGGCATTTGGCGGTATCA 1950  
Qy 6602 CAACTCTTACTCAGTTTCTACTACAAACTATCTCAAGAACTAAATCGGGACATGGAAACGG 6661  
Db 1951 CAACTCTTACTCAGTTTCTACTACAAACTATCTCAAGAACTAAATCGGGACATGGAAACGG 2010  
Qy 6662 TCGCGGACTCCCTGGTGCACCTTGGCAAGATCAACTTTAACTCCCTAGCAGCAGTAGTCCCTTC 6721  
Db 2011 TCGCGGACTCCCTGGTGCACCTTGGCAAGATCAACTTTAACTCCCTAGCAGCAGTAGTCCCTTC 2070

RESULT 21  
US-10-632-793-21  
; Sequence 21, Application US/10632793  
; Publication No. US20040048298A1  
; GENERAL INFORMATION:  
; APPLICANT: PARANHOS-BACCALA, Glaucia  
; APPLICANT: MALLET, Francois  
; APPLICANT: VOISSET, Cecile  
; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN  
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT  
; FILE REFERENCE: 110048  
; CURRENT APPLICATION NUMBER: US/10/632,793  
; CURRENT FILING DATE: 2003-08-04  
; PRIOR APPLICATION NUMBER: US/09/869,927  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: PCT/FR00/00144  
; PRIOR FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: FR 99/00888  
; PRIOR FILING DATE: 1999-01-21  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 21  
; LENGTH: 2938  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-632-793-21

Query Match 9.5%; Score 724; DB 8; Length 2938;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 724; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1476 GAATTACTCAATGATGTCACCATTAACACAGGGGAGGGAAGAAATCTCTGCTTTC 1535  
Db 883 GAATTACTCAATGATGTCACCATTAACACAGGGGAGGGAAGAAATCTCTGCTTTC 942

Qy 1536 TGGAGAGCTAAGGGAGGCAATTTGAGGAGCGCTCTCTGTCACCTGACTCTTCTGAAG 1595  
Db 943 TGGAGAGCTAAGGGAGGCAATTTGAGGAGCGCTCTCTGTCACCTGACTCTTCTGAAG 1002

Qy 1596 GCCAACTAATCTTAAAGCGTAAGTTTATCCTCAGTCAGCTCAGACATTAGAAAAA 1655  
Db 1003 GCCAACTAATCTTAAAGCGTAAGTTTATCCTCAGTCAGCTCAGACATTAGAAAAA 1062

Qy 1656 TCAAAAGTCTGCGTAGGCGGAGCAAACTTAGAAAACTTGAATTTGAACTTGGCACTC 1715  
Db 1063 TCAAAAGTCTGCGTAGGCGGAGCAAACTTAGAAAACTTGAATTTGAACTTGGCACTC 1122

Qy 1716 GGTTTTATATAGAGATCAGGAGCAGCGGAAACAGGACAAACGGGATTTAAAAAA 1775  
Db 1123 GGTTTTATATAGAGATCAGGAGCAGCGGAAACAGGACAAACGGGATTTAAAAAA 1182

Qy 1776 AGGCCAACCGCTTTAGTCATGACCTCAGGCAAGTGAAGCTTTGGAGGCTCTGAAAAAGGA 1835  
Db 1183 AGGCCAACCGCTTTAGTCATGACCTCAGGCAAGTGAAGCTTTGGAGGCTCTGAAAAAGGA 1242

Qy 1836 AAGCTGGGCAATTAAGTCCCTAATAGGCTTGTTCAGTGCCTCTACAGGACACT 1895  
Db 1243 AAGCTGGGCAATTAAGTCCCTAATAGGCTTGTTCAGTGCCTCTACAGGACACT 1302

Qy 1896 TTAATAAAGATTTGTCAAAGTAGAAGTAAGCGGCCCTCTGTCATGCCCTTTATTTCAAG 1955  
Db 1303 TTAATAAAGATTTGTCAAAGTAGAAGTAAGCGGCCCTCTGTCATGCCCTTTATTTCAAG 1362

Qy 1956 GGAATCACTGAAGSCCCTGCTGCCCGGACAAAGGTCTTTTGAATCAGAAGCCACTA 2015  
Db 1363 GGAATCACTGAAGSCCCTGCTGCCCGGACAAAGGTCTTTTGAATCAGAAGCCACTA 1422

Qy 2016 ACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCA 2075  
Db 1423 ACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCA 1482

Qy 2076 CCCTCACAGAGCCCTGGGTATGTTGACCATTTGAGGGCCAGGAAGGTTGTCTCTGGACA 2135  
Db 1483 CCCTCACAGAGCCCTGGGTATGTTGACCATTTGAGGGCCAGGAAGGTTGTCTCTGGACA 1542

Qy 2136 CTGGTGGGCTCTCTTACTGCTTCTCTGTCGCGGACAACTGTCCTCCAGATCTGTCA 2195  
Db 1543 CTGGTGGGCTCTCTTACTGCTTCTCTGTCGCGGACAACTGTCCTCCAGATCTGTCA 1602

Qy 2196 CTAT 2199  
Db 1603 CTAT 1606

RESULT 22  
US-10-717-580-2  
; Sequence 2, Application US/10717580  
; Publication No. US20040176314A1  
; GENERAL INFORMATION:  
; APPLICANT: BESEME, Frederic  
; APPLICANT: BLOUND, Jean-Luc  
; APPLICANT: BOUTON, Olivier  
; APPLICANT: MANDRAND, Bernard  
; APPLICANT: MALLET, Francois  
; APPLICANT: PERRON, Herve  
; TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISEASE  
; TITLE OF INVENTION: WITH PREGNANCY DISORDERS  
; FILE REFERENCE: 105045  
; CURRENT APPLICATION NUMBER: US/10/717,580  
; CURRENT FILING DATE: 2003-11-21  
; PRIOR APPLICATION NUMBER: US/09/446,024A  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: PCT/FR98/01442  
; PRIOR FILING DATE: 1998-07-06  
; PRIOR APPLICATION NUMBER: FR 97/08815  
; PRIOR FILING DATE: 1997-07-07  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 2938  
; TYPE: DNA  
; ORGANISM: Human  
US-10-717-580-2

Query Match 9.5%; Score 724; DB 9; Length 2938;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 724; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1476 GAATTACTCAATGATGTCACCATTAACACAGGGGAGGGAAGAAATCTCTGCTTTC 1535  
Db 883 GAATTACTCAATGATGTCACCATTAACACAGGGGAGGGAAGAAATCTCTGCTTTC 942

Qy 1536 TGGAGAGCTAAGGGAGGCAATTTGAGGAGCGCTCTCTGTCACCTGACTCTTCTGAAG 1595  
Db 943 TGGAGAGCTAAGGGAGGCAATTTGAGGAGCGCTCTCTGTCACCTGACTCTTCTGAAG 1002

Qy 1596 GCCAACTAATCTTAAAGCGTAAGTTTATCCTCAGTCAGCTCAGACATTAGAAAAA 1655  
Db 1003 GCCAACTAATCTTAAAGCGTAAGTTTATCCTCAGTCAGCTCAGACATTAGAAAAA 1062

Qy 1656 TCAAAAGTCTGCGTAGGCGGAGCAAACTTAGAAAACTTGAATTTGAACTTGGCACTC 1715  
Db 1063 TCAAAAGTCTGCGTAGGCGGAGCAAACTTAGAAAACTTGAATTTGAACTTGGCACTC 1122

Qy 1716 GGTTTTATATAGAGATCAGGAGCAGCGGAAACAGGACAAACGGGATTTAAAAAA 1775



[illegible]

```

RESULT 23
US-10-114-893-134
; Sequence 134, Application US/10114893
; Publication No. US20020193567A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Carlin-Duckett, McKeough
; APPLICANT: Kelleher, Kerry S.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: GI 6000-10A
; CURRENT APPLICATION NUMBER: US/10/114,893
; CURRENT FILING DATE: 2002-04-02
; EARLIER APPLICATION NUMBER: 09/413,232
; EARLIER FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 134
; LENGTH: 2946
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-114-893-134

Query Match          9.0%; Score 680; DB 6; Length 2946;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1030; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      5685  AGAGTTTCTATGAGAAATGCAGCGTCCCGAAATATTGATGCCCAATCGATAGGAGTCT 5744

```

1032	DB	AGAGTTTCTATGGGAAATGCGAGCGTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCT	1091
5745	QY	TTSTAAGGGAACCCCACTTTCACTGCGCCACACCCATATGTCGCCGAACTGCTATCACTC	5804
1092	DB	TTCTAAGGGAACCCCACTTTCACTGCGCCACCCCATATGTCGCCGAACTGCTATCACTC	1151
5805	QY	TGCCACTCTTTTGATGATGCAATACATCTATTTGAGACAGGAAAAAATGATTTAATCTCTAG	5864
1152	DB	TGCCACTCTTTTGCA TGCATGCAAA TACTCATTTATGTCAGACGAAAAATGATTTAATCTCTAG	1211
5865	QY	TTGTCTCTGGAGGACTTGGAGTCACTGCTCTGTGGACTTACTTCCACCCAACTGCTATGTC	5924
1212	DB	TTGTCTCTGGAGGACTTGGAGTCACTGCTCTGTGGACTTACTTCCACCCAACTGCTATGTC	1271
5925	QY	TGATGGGGGTGGAGTTCAGATCAGGCAAGAGAAAAAATGTTAAAAAGAGTAATCTCCCA	5984
1272	DB	TGATGGGGGTGGAGTTCAGATCAGGCAAGAGAAAAAATGTTAAAAAGAGTAATCTCCCA	1331
5985	QY	ACTCACCSGGGTACATGGCACTCTAGCCCTCAAGGACTAGATCTCTCAAAACTACAC	6044
1332	DB	ACTCACCSGGGTACATGGCACTCTAGCCCTCAAGGACTAGATCTCTCAAAACTACAC	1391
6045	QY	TGAACCCCTCCGTAACCATACTCGCCTGGTAAAGCCTATTTAATACCACCCCTCACTGGGCT	6104
1392	DB	TGAACCCCTCCGTACCCATACTCGCCTGGTAAAGCCTATTTAATACCACCCCTCACTGGGCT	1451
6105	QY	CCATGAGGTCTCGGCCCAAAACCTACTAACTGTTGGATATGCTCCCTGAACCTTCAR	6164
1452	DB	CCATGAGGTCTCGGCCCAAAACCTACTAACTGTTGGATATGCTCCCTGAACCTTCAG	1511
6165	QY	GCCATATGTTTCAATCCCTGACTCTGAACAAATGGAACTTCCAGCACAGAAATAAACAC	6224
1512	DB	GCCATATGTTTCAATCCCTGACTCTGAACAAATGGAACTTCCAGCACAGAAATAAACAC	1571
6225	QY	CAC TTCGGTTTTAGTAGGACCTCTGTTTCCAATSTGGAAATAACCCATCACTCAAACCT	6284
1572	DB	CAC TTCGGTTTTAGTAGGACCTCTGTTTCCAATSTGGAAATAACCCATCACTCAAACCT	1631
6285	QY	CACCTGTGTAAATTTAGCAATACTACATACAAACCACTCCCAATGCATCAGGTGGGT	6344
1632	DB	CACCTGTGTAAATTTAGCAATACTACATACAAACCACTCCCAATGCATCAGGTGGGT	1691
6345	QY	AACCTCTCCCAACAATAAGTCTGCCTACCCCTCAGGAATATTTTTGTCTGTGGTACCTC	6404
1692	DB	AACCTCTCCCAACAATAAGTCTGCCTACCCCTCAGGAATATTTTTGTCTGTGGTACCTC	1751
6405	QY	AGCCTATCGTGTGTAATGGCTCTTCAGAACTATATGCTGCTCCCTCTCATTTCTAGTGC	6464
1752	DB	AGCCTATCGTGTGTAATGGCTCTTCAGAACTATATGCTGCTCCCTCTCATTTCTAGTGC	1811
6465	QY	CCCVATGRCATCTACACTGAAACAAGATTTATACAGTTATGTCATATCTAAGCCCGCAA	6524
1812	DB	CCCVATGACCATCTACACTGAAACAAGATTTATACAAATTATGTCATATCTAAGCCCGCAA	1871
6525	QY	CAAAAGAGTACCAATCTTCTCTTTTGTATAGGAGCAGGAGTGTAGGTGCATAGGTAC	6584
1872	DB	CAAAAGAGTACCAATCTTCTCTTTTGTATAGGAGCAGGAGTGTAGGTGCATAGGTAC	1931
6585	QY	TGGCATTTGGGGTATCACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA	6644
1932	DB	TGGCATTTGGGGTATCACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA	1991
6645	QY	TGGGGACATGGAACGGGTGGCCGACTCCCTGGTCACTTTGCAAGATCAACTTTAATCCCT	6704
1992	DB	TGGGGACATGGAACGGGTGGCCGACTCCCTGGTCACTTTGCAAGATCAACTTTAATCCCT	2051
6705	QY	AGCAGCATGATGTCCTTC 6721	
2052	DB	AGCAGCATGATGTCCTTC 2068	

RESULT 24  
US-10-016-249-3

```
; Sequence 3, Application US/10016249
; Publication No. US200301000531
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Metberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Mi, Sha
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6006B.AJ172A
; CURRENT APPLICATION NUMBER: US/10/016,249
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: US/09/175,928
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2946
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-016-249-3

Query Match      9.0%; Score 680; DB 6; Length 2946;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1030; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGGAGAATGCAGCGTCCGCGAAATATTGATGCCCATCGTATAGGAGTCT 5744
Db 1032 AGAGTTTCTATGGAGAATGCAGCGTCCGCGAAATATTGATGCCCATCGTATAGGAGTCT 1091
QY 5745 TTSTAAGGGAACCCCACTTCACTGCCCAACCCATATGCCCGCAACTGCTATCACTC 5804
Db 1092 TTCTAAGGGAACCCCACTTCACTGCCCAACCCATATGCCCGCAACTGCTATCACTC 1151
QY 5805 TGCCACTCTTTGCATGCAATGCAAACTACTATTTATTTGGACAGAAATAATGATTAATCTTAG 5864
Db 1152 TGCCACTCTTTGCATGCAATGCAAACTACTATTTATTTGGACAGAAATAATGATTAATCTTAG 1211
QY 5865 TTGTCTCGAGACTTGGAGTCACTGTCTGTGGACTTACTTCCCAACCCCACTGATGTC 5924
Db 1212 TTGTCTCGAGACTTGGAGTCACTGTCTGTGGACTTACTTCCCAACCCCACTGATGTC 1271
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAGAAAGTAATCTCCA 5984
Db 1272 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAGAAAGTAATCTCCA 1331
QY 5985 ACTCACCGGGTACATGGCACTCTAGCCCTTACAAAGACTAGATCTCTCAAAACTACA 6044
Db 1332 ACTCACCGGGTACATGGCACTCTAGCCCTTACAAAGACTAGATCTCTCAAAACTACA 1391
QY 6045 TGAACCCCTCCGTACCCATCTCGCTGTAGCCCTTTATATACCCCTCACTGGGCT 6104
Db 1392 TGAACCCCTCCGTACCCATCTCGCTGTAGCCCTTTATATACCCCTCACTGGGCT 1451
QY 6105 CCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTTGGATATGCTTCCCTGAACTTCAR 6164
Db 1452 CCATGAGGTCTCGGCCCAAAACCCCTACTAACTGTTGGATATGCTTCCCTGAACTTCAG 1511
QY 6165 GCCATATGTTTCAATCCCTGTACTGTAACAATGGAACAACTTTAGACAGAAATAAACAC 6224
Db 1512 GCCATATGTTTCAATCCCTGTACTGTAACAATGGAACAACTTTAGACAGAAATAAACAC 1571
QY 6225 CACTTCCGTTTGTAGTAGGACCTCTGTTTCCATSTGGAAATAACCCATCTCAAACT 6284
Db 1572 CACTTCCGTTTGTAGTAGGACCTCTGTTTCCATCTGGAATAACCCATCTCAAACT 1631
QY 6285 CACCTGTGTAAATTTAGCAATACATACACCAACTCCCAATGATCAGGTGGGT 6344
Db 1632 CACCTGTGTAAATTTAGCAATACATACACCAACTCCCAATGATCAGGTGGGT 1691

; Sequence 543, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt SEQ_genes Version 1.0
; SEQ ID NO 543
; LENGTH: 3633
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-543

Query Match      8.6%; Score 651; DB 10; Length 3633;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGGAGAATGCAGCGTCCGCGAAATATTGATGCCCATCGTATAGGAGTCT 5744
Db 866 AGAGTTTCTATGGAGAATGCAGCGTCCGCGAAATATTGATGCCCATCGTATAGGAGTCT 925
QY 5745 TTSTAAGGGAACCCCACTTCACTGCCCAACCCATATGCCCGCAACTGCTATCACTC 5804
Db 926 TTSTAAGGGAACCCCACTTCACTGCCCAACCCATATGCCCGCAACTGCTATCACTC 985
QY 5805 TGCCACTCTTTGCATGCAATGCAAACTACTATTTATTTGGACAGAAATAATGATTAATCTTAG 5864
Db 986 TGCCACTCTTTGCATGCAATGCAAACTACTATTTATTTGGACAGAAATAATGATTAATCTTAG 1045
QY 5865 TTGTCTCGAGACTTGGAGTCACTGTCTGTGGACTTACTTCCCAACCCCACTGATGTC 5924
Db 1046 TTGTCTCGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCCCAACCCCACTGATGTC 1105

RESULT 25
US-10-821-234-543
; Sequence 543, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt SEQ_genes Version 1.0
; SEQ ID NO 543
; LENGTH: 3633
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-543
```

```
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAACATGTAAAGAGTAATCTCCCA 5984
Db 1106 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAACATGTAAAGAGTAATCTCCCA 1165
QY 5985 ACTCAACCGGGTACATGGCACTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACA 6044
Db 1166 ACTCAACCGGGTACATGGCACTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACA 1225
QY 6045 TGAACCTCTCGTTACCATCTCTCGCTGGTAAAGCCTATTTAATAACCAACCTCACTGGGCT 6104
Db 1226 TGAACCTCTCGTTACCATCTCTCGCTGGTAAAGCCTATTTAATAACCAACCTCACTGGGCT 1285
QY 6105 CCATGAGTCTCGGCCCAAAACCTTACTACTGTTGGATATGCTCCCTCGGACTTCAR 6164
Db 1286 CCATGAGTCTCGGCCCAAAACCTTACTACTGTTGGATATGCTCCCTCGGACTTCAG 1345
QY 6165 GCCATATGTTTCAATCCCTGTACTGAACAATGGAACAACCTTCAGCACAGAAATAAACA 6224
Db 1346 GCCATATGTTTCAATCCCTGTACTGAACAATGGAACAACCTTCAGCACAGAAATAAACA 1405
QY 6225 CACTTCGGTTTGTAGTACGACTCTGTTTTCATSTGGAATAAACCCATACCTCAAACT 6284
Db 1406 CACTTCGGTTTGTAGTACGACTCTGTTTTCATSTGGAATAAACCCATACCTCAAACT 1465
QY 6285 CACTGTTGTAATTTAGCAATACTACATACACAAACCACTCCCAATGATCAGTGGGT 6344
Db 1466 CACTGTTGTAATTTAGCAATACTACATACACAAACCACTCCCAATGATCAGTGGGT 1525
QY 6345 AACTCTCCCAACAAATAGTCTGCCCTACCTCAGGAATAATTTTCTCTGTGTACTCTC 6404
Db 1526 AACTCTCCCAACAAATAGTCTGCCCTACCTCAGGAATAATTTTCTCTGTGTACTCTC 1585
QY 6405 AGCTATGCTGTTGTAATGGCTCTTCAGAAATCTATGTGCTTCTCTCATTTCTAGTGCC 6464
Db 1586 AGCTATGCTGTTGTAATGGCTCTTCAGAAATCTATGTGCTTCTCTCATTTCTAGTGCC 1645
QY 6465 CCCATGRCATCTACACTGAACAGATTTATACAGTTATGTCATATCTAAGCCCGCAA 6524
Db 1646 CCTATGACCATCTACACTGAACAGATTTATACAGTTATGTCATATCTAAGCCCGCAA 1705
QY 6525 CAAAAGATACCCATCTCTCTTTTGTATAGGAGCAGGAGTGTAGGTGCATCAGTAC 6584
Db 1706 CAAAAGATACCCATCTCTCTTTTGTATAGGAGCAGGAGTGTAGGTGCATCAGTAC 1765
QY 6585 TGCATTGGCGGTATCAACACCTCTACTCAGTTTCTACTACAAACTATCTCAAGAACTAAA 6644
Db 1766 TGCATTGGCGGTATCAACACCTCTACTCAGTTTCTACTACAAACTATCTCAAGAACTAAA 1825
QY 6645 TGGGACATGGAAGGGTCCCGACTCCCTGGTCACTTCCCAAGATCA 6692
Db 1826 TGGGACATGGAAGGGTCCCGACTCCCTGGTCACTTCCCAAGATCA 1873
```

## RESULT 26

```
US-10-632-793-23
; Sequence 23, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glauclia
; APPLICANT: MALLET, Françoise
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
; FILE REFERENCE: 110048
; CURRENT APPLICATION NUMBER: US/10/632,793
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/FR00/00144
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
```

## RESULT 27

```
US-10-717-580-4
; Sequence 4, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
; APPLICANT: BESEME, Frederic
; APPLICANT: BLOND, Jean-Luc
; APPLICANT: BOUTON, Olivier
; APPLICANT: MANDRAND, Bernard
; APPLICANT: MALLET, Francois
; APPLICANT: PERON, Herve
; TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DIS
; TITLE OF INVENTION: WITH PREGNANCY DISORDERS
; FILE REFERENCE: 105045
; CURRENT APPLICATION NUMBER: US/10/717,580
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US/09/446,024A
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/FR98/01442
; PRIOR FILING DATE: 1998-07-06
; NUMBER OF SEQ ID NOS: 1122
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 2006
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (305)..(305)
; OTHER INFORMATION: n = a or g or c or t/u
US-10-632-793-23
```

```
Query Match 5.3%; Score 399; DB 8; Length 2006;
Best Local Similarity 99.8%; Pred. No. 8.5e-201;
Matches 499; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4802 AGSCCTTCCCTCTAGGACAGAAAAGCCCAAGAGGTAAATAAGGCACTAGTTTCATGAAA 4861
Db 643 AGSCCTTCCCTCTAGGACAGAAAAGCCCAAGAGGTAAATAAGGCACTAGTTTCATGAAA 702
QY 4862 TAATTCAGATTCGGAATTCCTCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTCCAGG 4921
Db 703 TAATTCAGATTCGGAATTCCTCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTCCAGG 762
QY 4922 CCACAGTACCCAGGAGATATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCT 4981
Db 763 CCACAGTAAACCCAGGAGATATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCT 822
QY 4982 GAAGGCCACAGTCTCAGGGAAGTTCGAGAAAATGAATGAATACTCAAGGACATCTAA 5041
Db 823 GAAGGCCACAGTCTCAGGGAAGTTCGAGAAAATGAATGAATACTCAAGGACATCTAA 882
QY 5042 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCTGTCTGTGGCTTATAGCCCTTAAAAA 5101
Db 883 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCTGTCTGTGGCTTATAGCCCTTAAAAA 942
QY 5102 GAATCTGCAACTTCCCAAAAAGAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 5161
Db 943 GAATCTGCAACTTCCCAAAAAGAGGACTTAGCCCATACGAAATGCTGTATGGAAGGC 1002
QY 5162 CCTTCATACCAATGACCTTGTGCTGACCCAGAGCAGCACTTGTAGTTCACACATCAC 5221
Db 1003 CCTTCATACCAATGACCTTGTGCTGACCCAGAGCAGCACTTGTAGTTCACACATCAC 1062
QY 5222 CTCCTTAGCCAAATATCAACAAGTTCTTAAACATTACAGGAACCTATCCCTGAGAAGA 5281
Db 1063 CTCCTTAGCCAAATATCAACAAGTTCTTAAACATTACAGGAACCTATCCCTGAGAAGA 1122
QY 5282 GGGAAAAGAACTATTCACCC 5302
Db 1123 GGGAAAAGAACTATTCACCC 1143
```

; PRIOR FILING DATE: 1997-07-07  
 ; NUMBER OF SEQ ID NOS: 53  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 4  
 ; LENGTH: 2006  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (305)..(305)  
 ; OTHER INFORMATION: n = any nucleotide  
 US-10-717-580-4

Query Match 5.3%; Score 399; DB 9; Length 2006;  
 Best Local Similarity 99.6%; Pred. No. 8.5e-201;  
 Matches 499; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 4802 AGGCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTATAAAGGCATAGTTCATGAAA 4861  
 DB 643 AGGCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTATAAAGGCATAGTTCATGAAA 702  
 QY 4862 TAATTCACAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTCCAGG 4921  
 DB 703 TAATTCACAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTCCAGG 762  
 QY 4922 CCACAGTAACCCAGGAGGTATCCAGGCGTTAGGTATACGATATCACTTACACTGGCCT 4981  
 DB 763 CCACAGTAACCCAGGAGGTATCCAGGCGTTAGGTATACGATATCACTTACACTGGCCT 822  
 QY 4982 GAAGGCCACAGTCTCCAGGGAAGGTCGAGAAAATGAATGAAAVACTCAAGAGGACATCTAA 5041  
 DB 823 GAAGGCCACAGTCTCCAGGGAAGGTCGAGAAAATGAATGAAAVACTCAAGAGGACATCTAA 882  
 QY 5042 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTGGCTATAGCTTAAAAA 5101  
 DB 883 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTGGCTATAGCTTAAAAA 942  
 QY 5102 GAATCTGCAACTTCCGAAAAGAGGAGCTTACGCAATACGAAATGCTGTATGAGGC 5161  
 DB 943 GAATCTGCAACTTCCGAAAAGAGGAGCTTACGCAATACGAAATGCTGTATGAGGC 1002  
 QY 5162 CTTTCAATACCAATGACCTTGTGCTTACCAAGAGGAGGAGCTTGTAGTTCAGAGATCAC 5221  
 DB 1003 CTTTCAATACCAATGACCTTGTGCTTACCAAGAGGAGGAGCTTGTAGTTCAGAGATCAC 1062  
 QY 5222 CTCCTTAGCCAAATATCAACAAAGTTCTTAAACATTTACAAAGAACTATCCCTGAGAAGA 5281  
 DB 1063 CTCCTTAGCCAAATATCAACAAAGTTCTTAAACATTTACAAAGAACTATCCCTGAGAAGA 1122  
 QY 5282 GGGAAAAGAACTATTCACCC 5302  
 DB 1123 GGGAAAAGAACTATTCACCC 1143

RESULT 28  
 US-10-450-763-20013  
 ; Sequence 20013, Application US/10450763  
 ; Publication No. US20050196754A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hyseq, Inc  
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
 ; FILE REFERENCE: 790CIP3/US  
 ; CURRENT APPLICATION NUMBER: US/10/450,763  
 ; CURRENT FILING DATE: 2003-06-11  
 ; PRIOR APPLICATION NUMBER: PCT/US01/08631  
 ; PRIOR FILING DATE: 2001-03-30  
 ; PRIOR APPLICATION NUMBER: 09/540,217  
 ; PRIOR FILING DATE: 2000-03-31  
 ; PRIOR APPLICATION NUMBER: 09/649,167  
 ; PRIOR FILING DATE: 2000-08-23  
 ; NUMBER OF SEQ ID NOS: 60736  
 ; SOFTWARE: Custom  
 ; SEQ ID NO 20013

; LENGTH: 8294  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SIMILAR  
 ; LOCATION: (1153)..(2754)  
 ; OTHER INFORMATION: 89% homologous to Homo sapiens PEX1, accession number  
 ; OTHER INFORMATION: AB008112, Smith-Waterman Score=2320.  
 US-10-450-763-20013

Query Match 5.3%; Score 399; DB 10; Length 8294;  
 Best Local Similarity 99.6%; Pred. No. 8.6e-201;  
 Matches 499; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 4802 AGGCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTATAAAGGCATAGTTCATGAAA 4861  
 DB 6964 AGGCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTATAAAGGCATAGTTCATGAAA 7023  
 QY 4862 TAATTCACAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTCCAGG 4921  
 DB 7024 TAATTCACAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTCCAGG 7083  
 QY 4922 CCACAGTAACCCAGGAGGTATCCAGGCGTTAGGTATACGATATCACTTACACTGGCCT 4981  
 DB 7084 CCACAGTAACCCAGGAGGTATCCAGGCGTTAGGTATACGATATCACTTACACTGGCCT 7143  
 QY 4982 GAAGGCCACAGTCTCCAGGGAAGGTCGAGAAAATGAATGAAAVACTCAAGAGGACATCTAA 5041  
 DB 7144 GAAGGCCACAGTCTCCAGGGAAGGTCGAGAAAATGAATGAAAVACTCAAGAGGACATCTAA 7203  
 QY 5042 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTGGCTATAGCTTAAAAA 5101  
 DB 7204 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCCTGYTCTGTGGCTATAGCTTAAAAA 7263  
 QY 5102 GAATCTGCAACTTCCGAAAAGAGGAGCTTACGCAATACGAAATGCTGTATGAGGC 5161  
 DB 7264 GAATCTGCAACTTCCGAAAAGAGGAGCTTACGCAATACGAAATGCTGTATGAGGC 7323  
 QY 5162 CTTTCAATACCAATGACCTTGTGCTTACCAAGAGGAGGAGCTTGTAGTTCAGAGATCAC 5221  
 DB 7324 CTTTCAATACCAATGACCTTGTGCTTACCAAGAGGAGGAGCTTGTAGTTCAGAGATCAC 7383  
 QY 5222 CTCCTTAGCCAAATATCAACAAAGTTCTTAAACATTTACAAAGAACTATCCCTGAGAAGA 5281  
 DB 7384 CTCCTTAGCCAAATATCAACAAAGTTCTTAAACATTTACAAAGAACTATCCCTGAGAAGA 7443  
 QY 5282 GGGAAAAGAACTATTCACCC 5302  
 DB 7444 GGGAAAAGAACTATTCACCC 7464

RESULT 29  
 US-10-632-793-28  
 ; Sequence 28, Application US/10632793  
 ; Publication No. US20040048298A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: PARANHOS-BACCALA, Glauca  
 ; APPLICANT: MALLET, Francois  
 ; APPLICANT: VOISSET, Cecile  
 ; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN  
 ; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT  
 ; FILE REFERENCE: 110048  
 ; CURRENT APPLICATION NUMBER: US/10/632,793  
 ; CURRENT FILING DATE: 2003-08-04  
 ; PRIOR APPLICATION NUMBER: US/09/869,927  
 ; PRIOR FILING DATE: 2001-10-22  
 ; PRIOR APPLICATION NUMBER: PCT/FR00/00144  
 ; PRIOR FILING DATE: 2000-01-21  
 ; PRIOR APPLICATION NUMBER: FR 99/00888  
 ; PRIOR FILING DATE: 1999-01-21  
 ; NUMBER OF SEQ ID NOS: 33  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 28

```
; LENGTH: 3372
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-632-793-28

Query Match      5.0%; Score 380; DB 8; Length 3372;
Best Local Similarity 99.6%; Pred. No. 1.2e-190;
Matches 480; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4802 AGGCTTCCCTGTAGACAGAGAAAGGCCCAAGAGGTAATAAAGGCACTAGTTTCATGAAA 4861
Db 2032 AGGCTTCCCTGTAGACAGAGAAAGGCCCAAGAGGTAATAAAGGCACTAGTTTCATGAAA 2091
QY 4862 TAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACATAGCCCTGCTTTCAGG 4921
Db 2092 TAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACATAGCCCTGCTTTCAGG 2151
QY 4922 CCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTTGGCCCT 4981
Db 2152 CCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTTGGCCCT 2211
QY 4982 GAAGGCCACAGTCTCAGGAGAGGTTCGAGAAAATGAATGAAAYACTCAAAAGGACATCTAA 5041
Db 2212 GAAGGCCACAGTCTCAGGAGAGGTTCGAGAAAATGAATGAAAYACTCAAAAGGACATCTAA 2271
QY 5042 AAAGCAAAACCCAGGAAACCCACTCACATGGGCTGTCTGTGGCTTATAGCCTTAAAAA 5101
Db 2272 AAAGCAAAACCCAGGAAACCCACTCACATGGGCTGTCTGTGGCTTATAGCCTTAAAAA 2331
QY 5102 GAATCTGCAACTTTCCTCAAAAGCAGGACTTAGCCCATACGAAATGCTCTATGGAGGC 5161
Db 2332 GAATCTGCAACTTTCCTCAAAAGCAGGACTTAGCCCATACGAAATGCTCTATGGAGGC 2391
QY 5162 CCTTCATAACCAATGACCTTGTGTGACCCAAAGACAGCCAACTTAGTTGCAGACATCAC 5221
Db 2392 CCTTCATAACCAATGACCTTGTGTGACCCAAAGACAGCCAACTTAGTTGCAGACATCAC 2451
QY 5222 CTCCTTAGCCAAATATCAAGTTCCTTAAAAACATTACAAGGAACCTTATCCCTGAGAAGA 5281
Db 2452 CTCCTTAGCCAAATATCAAGTTCCTTAAAAACATTACAAGGAACCTTATCCCTGAGAAGA 5281
QY 5282 GG 5283
Db 2512 GG 2513

Search completed: August 2, 2006, 15:51:59
Job time : 7491 secs
```

```
RESULT 30
US-10-717-580-9
; Sequence 9, Application US/10717580
; Publication No. US20040176314A1
; GENERAL INFORMATION:
; APPLICANT: BESEME, Frederic
; APPLICANT: BLOND, Jean-Luc
; APPLICANT: BOUTON, Olivier
; APPLICANT: MANDRAND, Bernard
; APPLICANT: MALLET, Francois
; APPLICANT: PERRON, Herve
; TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DISEASES
; TITLE OF INVENTION: WITH PREGNANCY DISORDERS
; FILE REFERENCE: 105045
; CURRENT APPLICATION NUMBER: US/10717,580
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US/09/446,024A
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/FR98/01442
; PRIOR FILING DATE: 1998-07-06
; PRIOR APPLICATION NUMBER: FR 97/08815
; PRIOR FILING DATE: 1997-07-07
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 3372
; TYPE: DNA
```

**This Page Blank (uspto)**



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2006, 09:10:38.; Search time 4018 Seconds  
(without alignments)  
13156.692 Million cell updates/sec

Title: US-10-717-580-11  
Perfect score: 7582  
Sequence: 1 caacaatcggtataataacc.....tattaaatctgcactgtgc 7582

Scoring table: OLIGO\_NUC  
Gapop 60.0, Capext 60.0

Searched: 5244920 seqs, 3486124231 residues

Word size : 6

Total number of hits satisfying chosen parameters: 10438277

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 90 summaries

Database : N Geneseq 8.\*

- 1: geneseqn1980s.\*
- 2: geneseqn1990s.\*
- 3: geneseqn2000s.\*
- 4: geneseqn2001as.\*
- 5: geneseqn2001bs.\*
- 6: geneseqn2002as.\*
- 7: geneseqn2002bs.\*
- 8: geneseqn2003as.\*
- 9: geneseqn2003bs.\*
- 10: geneseqn2003cs.\*
- 11: geneseqn2003ds.\*
- 12: geneseqn2004as.\*
- 13: geneseqn2004bs.\*
- 14: geneseqn2005s.\*
- 15: geneseqn2006s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7582	100.0	7582	3	Aax59215 Human end
2	7378	97.3	7582	2	Aax25665 Complete
3	979	12.9	7466	5	Aax68626 DNA encod
4	888	11.7	2372	2	Aax25664 Human end
5	888	11.7	2372	3	Aax59214
6	779	10.3	2599	3	Abn97927 Human ret
7	779	10.3	8523	7	Ady30988 Human gen
8	779	10.3	8523	7	Ady36376 HIRA geno
9	779	10.3	10499	3	Abn97929 Human ret
10	779	10.3	56093	6	Abn161744
11	750	9.9	6394	5	Aax84210 DNA encod
12	731	9.6	1614	14	AED10984
13	731	9.6	1614	14	AED10986
14	731	9.6	1617	5	Aax20070 HERV-W en
15	731	9.6	1617	14	AED10942
16	731	9.6	1948	2	Aax25659 Human end
17	731	9.6	1948	5	Aax59209 5' non co
18	731	9.6	2055	3	Abn97948 Human ret

## ALIGNMENTS

```
RESULT 1
AAAS9215
ID AAAS9215 standard; DNA; 7582 BP.
XX
AC AAAS9215;
XX
DT 07-NOV-2000 (first entry)
XX
DE Human endogenous retrovirus W (HERV-W) sequence.
XX
KW Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
KW gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
XX
OS Human endogenous retrovirus.
XX
FH Key Location/Qualifiers
FT LTR 1..120
FT /*tag= a
FT /note= "R of 5' LTR"
FT LTR 121..575
FT /*tag= b
FT /*tag= "U5 of 5' LTR"
FT primer_bind 579..596
FT /*tag= c
FT CDS 581..7194
FT /*tag= d
FT /*tag= "ORF1 env538"
FT CDS 7039..7194
FT /*tag= e
FT /*tag= "ORF2 52 AA"
FT CDS 7112..7255
FT /*tag= f
FT /*tag= "ORF3 48 AA"
FT misc_feature 7244..7254
FT /*tag= g
FT /note= "polypurine tract"
FT LTR 7256..7582
FT /*tag= h
FT polyA_signal 7563..7569
FT /*tag= i
XX
WO200043521-A2.
XX
PD 27-JUL-2000.
XX
PF 21-JAN-2000; 2000WO-FR000144.
XX
PR 21-JAN-1999; 99FR-00000888.
XX
PA (INNR ) BIO MERIEUX.
XX
PI Paranhos-Baccala G, Mallet F, Voisset C;
XX
DR WPI; 2000-499229/44.
XX
PT New nucleic acid from human endogenous retrovirus, useful e.g. for
PT diagnosis of autoimmune disease and complications of pregnancy, contains
PT at least part of the gag gene.
XX
PS Disclosure; Page 49-52; 53pp; French.
XX
CC The present sequence represents an endogenous retrovirus, which is
CC associated with an autoimmune disease, and is integrated into the human
CC genome. The retrovirus is human endogenous retrovirus W (HERV-W). The
CC HERV-W retrovirus is associated with autoimmune disease, failure of
CC pregnancy or disorders of pregnancy. HERV-W nucleic acid fragments, or
CC proteins derived from it, are useful for diagnosis of autoimmune disease
CC (specifically multiple sclerosis) and for monitoring pregnancy. The
```

CC nucleic acid fragments may also be used for in situ labelling of isolated  
CC chromosomes, while the transcription product can be used to study or  
CC monitor T cell proliferation in vitro

```
SQ Sequence 7582 BP; 2156 A; 1876 C; 1538 G; 1796 T; 0 U; 216 Other;
Query Match 100.0%; Score 7582; DB 3; Length 7582;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CAACAATCGGGATATAAAACCCAGGCATTTCGAGCTGGCAACAGACGCCCTTTGGGTCC 60
Db |||||||
Qy 1 CAACAATCGGGATATAAAACCCAGGCATTTCGAGCTGGCAACAGACGCCCTTTGGGTCC 60
Db |||||||
Qy 61 CTTCCCTTTGATGGGAGCTGTTTTCATGCTATTTCACCTCTATTAAATCTTGCACACTGA 120
Db |||||||
Qy 61 CTTCCCTTTGATGGGAGCTGTTTTCATGCTATTTCACCTCTATTAAATCTTGCACACTGA 120
Db |||||||
Qy 121 CTCTTCTGGTCCATGTTTCTTACGCTCGAGCTGAGCTTTTGTCTCACCCTCCACCACATGC 180
Db |||||||
Qy 121 CTCTTCTGGTCCATGTTTCTTACGCTCGAGCTGAGCTTTTGTCTCACCCTCCACCACATGC 180
Db |||||||
Qy 181 TGTTCACACACCGCANACCTGCGCTGACTCCCATCCCTCTGATCCTGACGGGTGC 240
Db |||||||
Qy 181 TGTTCACACACCGCANACCTGCGCTGACTCCCATCCCTCTGATCCTGACGGGTGC 240
Db |||||||
Qy 241 CGCTGTGCTCTGATCCAGCGAGCGGCCCATTTGGCGTCCCAATTTGGGCTAAAGCTTGC 300
Db |||||||
Qy 241 CGCTGTGCTCTGATCCAGCGAGCGGCCCATTTGGCGTCCCAATTTGGGCTAAAGCTTGC 300
Db |||||||
Qy 301 CATTTGNCCTGCACGGCTAAGTGCTGGGTTTGTCTAAATTGAGCTGAACACTANTCACT 360
Db |||||||
Qy 301 CATTTGNCCTGCACGGCTAAGTGCTGGGTTTGTCTAAATTGAGCTGAACACTANTCACT 360
Db |||||||
Qy 361 GGGTTCATGTTCTCTTCTGACCCACGGCTTCTAATAKAACTATATAACACTTACACACA 420
Db |||||||
Qy 361 GGGTTCATGTTCTCTTCTGACCCACGGCTTCTAATAKAACTATATAACACTTACACACA 420
Db |||||||
Qy 421 TGGCCCAAGATTCCATTCTTGGAAATCGTGAGGSAACGAACTCCAGGTGAGAAATAC 480
Db |||||||
Qy 421 TGGCCCAAGATTCCATTCTTGGAAATCGTGAGGSAACGAACTCCAGGTGAGAAATAC 480
Db |||||||
Qy 481 GARGCTTGCCACCATCTTGGAAAGCGGCTGTACCTCTCGAAAGTGTTTACACCACTC 540
Db |||||||
Qy 481 GARGCTTGCCACCATCTTGGAAAGCGGCTGTACCTCTCGAAAGTGTTTACACCACTC 540
Db |||||||
Qy 541 TTGGGAGCTCTGTGAGCAAGGACCCCGGCTACATTTTGGCRACCAACGAGCATCC 600
Db |||||||
Qy 541 TTGGGAGCTCTGTGAGCAAGGACCCCGGCTACATTTTGGCRACCAACGAGCATCC 600
Db |||||||
Qy 601 MAAGTGATGGGAAACGTTCCCGCAAGCAAAAAGCCCTTAAGACGTATTTCGGARAAT 660
Db |||||||
Qy 601 MAAGTGATGGGAAACGTTCCCGCAAGCAAAAAGCCCTTAAGACGTATTTCGGARAAT 660
Db |||||||
Qy 661 TGGGACCAATTTGACCTCAGACACTAAGAAGAAACGACTTATATTCTCTGCAAGTCC 720
Db |||||||
Qy 661 TGGGACCAATTTGACCTCAGACACTAAGAAGAAACGACTTATATTCTCTGCAAGTCC 720
Db |||||||
Qy 721 GCCTGGCACTCTCTGAGGGAAGTATAAATTATAACCACTTTTACAGCTAGACTCTTTTG 780
Db |||||||
Qy 721 GCCTGGCACTCTCTGAGGGAAGTATAAATTATAACCACTTTTACAGCTAGACTCTTTTG 780
Db |||||||
Qy 781 TAGAAAAGGCAAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTTCATTAAAGACAA 840
Db |||||||
Qy 781 TAGAAAAGGCAAAATGGAGTGAAGTGCCATAAGTACAAACTTTCTTTTTCATTAAAGACAA 840
Db |||||||
Qy 841 CTCACAATTATGTAAGAGTGTATTTATGCGCTTACAGAGCCCTTACAGCTTACCTCC 900
Db |||||||
Qy 841 CTCACAATTATGTAAGAGTGTATTTATGCGCTTACAGAGCCCTTACAGCTTACCTCC 900
Db |||||||
Qy 901 CTATCCCAAGCATCCCGACTCTCTTCCCAATTAAGGACCCCTTTCAACCCAAATGG 960
Db |||||||
Qy 901 CTATCCCAAGCATCCCGACTCTCTTCCCAATTAAGGACCCCTTTCAACCCAAATGG 960
Db |||||||
```

QY 961 TCCAAAAGGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTGGCCAAATATTTCCCAAT 1020  
DB 961 TCCAAAAGGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTGGCCAAATATTTCCCAAT 1020  
QY 1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAAATTCGGCCCGAGCCAGAGTGCATGTGCT 1080  
DB 1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAAATTCGGCCCGAGCCAGAGTGCATGTGCT 1080  
QY 1081 TTTTYYTCTCCGAGCTTAAAGCAAAATAAAACAGACTTAGGTAAATTTCTCAGATAAYCCT 1140  
DB 1081 TTTTYYTCTCCGAGCTTAAAGCAAAATAAAACAGACTTAGGTAAATTTCTCAGATAAYCCT 1140  
QY 1141 GATGGCTATATGRTGTTTTACAGGGTTAGGACAAATTTCTGATCTGACATGGAGAGAT 1200  
DB 1141 GATGGCTATATGRTGTTTTACAGGGTTAGGACAAATTTCTGATCTGACATGGAGAGAT 1200  
QY 1201 ATATATGTCACTGCTAAATCAGACACTAACCACCAATGAGAGAAGTGCCACCACTAACTGC 1260  
DB 1201 ATATATGTCACTGCTAAATCAGACACTAACCACCAATGAGAGAAGTGCCACCACTAACTGC 1260  
QY 1261 AGCTGAGRGTTGGGATCTCTGGTATCTCAGTCAAGTCAATGGATPANGATGACAAACA 1320  
DB 1261 AGCTGAGRGTTGGGATCTCTGGTATCTCAGTCAAGTCAATGGATPANGATGACAAACA 1320  
QY 1321 GAAGGAAGANATGATTTCCCAAGCCAGCCAGCAGCAGTTCCAGTCTASACCTCATTTG 1380  
DB 1321 GAAGGAAGANATGATTTCCCAAGCCAGCCAGCAGCAGTTCCAGTCTASACCTCATTTG 1380  
QY 1381 GGGACACAGAAATCAGTAACATGGGATTTGGTGTGTCAGACATTTTCTAACTTTGTGTC 1440  
DB 1381 GGGACACAGAAATCAGTAACATGGGATTTGGTGTGTCAGACATTTTCTAACTTTGTGTC 1440  
QY 1441 TASAAAGACTTAAGGAAAACCTASGAAGAAATCTAYGAAATTTACTCAATGATGTCCACCATA 1500  
DB 1441 TASAAAGACTTAAGGAAAACCTASGAAGAAATCTAYGAAATTTACTCAATGATGTCCACCATA 1500  
QY 1501 ACACAGGGNAGGAAGAAATCTTACTGCTTTCTGGAGAGACTAAGGAGGCAATTTGAG 1560  
DB 1501 ACACAGGGNAGGAAGAAATCTTACTGCTTTCTGGAGAGACTAAGGAGGCAATTTGAG 1560  
QY 1561 GAAGCGTGCTCTGTGCTCACTGCTCTCTGAAGGCCCACTAATCTTAAAGCGTAAGTT 1620  
DB 1561 GAAGCGTGCTCTGTGCTCACTGCTCTCTGAAGGCCCACTAATCTTAAAGCGTAAGTT 1620  
QY 1621 TATCACTCAGTCACTGTCAGACATTAAGAAAACCTTCAAAAGTCTGCGGTAGGCCCGGAG 1680  
DB 1621 TATCACTCAGTCACTGTCAGACATTAAGAAAACCTTCAAAAGTCTGCGGTAGGCCCGGAG 1680  
QY 1681 CAAAACCTTAGAAACCTTAGAACTTGGCAACVTCGGTTTTTATAATAGAGATCAGGAG 1740  
DB 1681 CAAAACCTTAGAAACCTTAGAACTTGGCAACVTCGGTTTTTATAATAGAGATCAGGAG 1740  
QY 1741 GAGCAGCGGAACAGACAAACGGGATTAAGAAAAGGCCACCGCTTTAGTCATGACCCT 1800  
DB 1741 GAGCAGCGGAACAGACAAACGGGATTAAGAAAAGGCCACCGCTTTAGTCATGACCCT 1800  
QY 1801 CAGGCAAGTGCACTTTGGAGGCTCTGGAAGGGAAAAGCTGGGCAAAATTTGAATGCTTAA 1860  
DB 1801 CAGGCAAGTGCACTTTGGAGGCTCTGGAAGGGAAAAGCTGGGCAAAATTTGAATGCTTAA 1860  
QY 1861 TAGGGCTTGCTTCCAGTGGGCTTACAAAGGACACTTTAAAAAGATTTGTCCAAGTAGAAG 1920  
DB 1861 TAGGGCTTGCTTCCAGTGGGCTTACAAAGGACACTTTAAAAAGATTTGTCCAAGTAGAAG 1920  
QY 1921 TAAGCGCCCTCTGCTCATGCCCTTAATTTCAAGGGAATCACTGGAAGGCCACCTGCCC 1980  
DB 1921 TAAGCGCCCTCTGCTCATGCCCTTAATTTCAAGGGAATCACTGGAAGGCCACCTGCCC 1980  
QY 1981 CAGGGACAAAGGCTTTTGGTTCAGAGCCACTAACAGATGATCCAGCAGGAGACTG 2040  
DB 1981 CAGGGACAAAGGCTTTTGGTTCAGAGCCACTAACAGATGATCCAGCAGGAGACTG 2040

QY 2041 AGGGTGCTGGGCAAGCGCCATCCCATGCCATCACTCATCAGAGCCCTGGGTATGCTT 2100  
DB 2041 AGGGTGCTGGGCAAGCGCCATCCCATGCCATCACTCATCAGAGCCCTGGGTATGCTT 2100  
QY 2101 GACCATTTAGGGGCCAGAAAGTTGTCTCTGGACACTGGTGCGGTCTTCTTAGTCTTACT 2160  
DB 2101 GACCATTTAGGGGCCAGAAAGTTGTCTCTGGACACTGGTGCGGTCTTCTTAGTCTTACT 2160  
QY 2161 CTTCTGTCCCGGCAACTGTCTCTCAGATCTGTCACTATTTCTGAGGGGTCCNTAAGACG 2220  
DB 2161 CTTCTGTCCCGGCAACTGTCTCTCAGATCTGTCACTATTTCTGAGGGGTCCNTAAGACG 2220  
QY 2221 GGCAGTCACTAGATATCTTTTCCAGCCACTAAGTTATGAACTGGGAGGCTTTATTTCT 2280  
DB 2221 GGCAGTCACTAGATATCTTTTCCAGCCACTAAGTTATGAACTGGGAGGCTTTATTTCT 2280  
QY 2281 TCACATGCTTTTCTAATTTATGTTGAAAGCCCACTACCTTGTAGGAGAGACATTTCTA 2340  
DB 2281 TCACATGCTTTTCTAATTTATGTTGAAAGCCCACTACCTTGTAGGAGAGACATTTCTA 2340  
QY 2341 GCAAAAGCAGGGGCCATTATACCTGAAACATAGGAGAAGGAAACCCCGTTTGTGTNCC 2400  
DB 2341 GCAAAAGCAGGGGCCATTATACCTGAAACATAGGAGAAGGAAACCCCGTTTGTGTNCC 2400  
QY 2401 CCTGCTTGAGGAAGGAATTAATCTGAACTCTGGGCAACAGAAAGGACATATATGACGAGC 2460  
DB 2401 CCTGCTTGAGGAAGGAATTAATCTGAACTCTGGGCAACAGAAAGGACATATATGACGAGC 2460  
QY 2461 CAAAGAAATCCCGTCTGTTCAAGTTAAACTAAGGATTTCCACTTCTTCCCTACCAAA 2520  
DB 2461 CAAAGAAATCCCGTCTGTTCAAGTTAAACTAAGGATTTCCACTTCTTCCCTACCAAA 2520  
QY 2521 GGCAGTACCCCTCAGACCCCAAGGCCCAACAGGATTTCAAAGATTTGTTAAGGACTTAA 2580  
DB 2521 GGCAGTACCCCTCAGACCCCAAGGCCCAACAGGATTTCAAAGATTTGTTAAGGACTTAA 2580  
QY 2581 AAGCCCAAGGCTTTAGTAAACCAATGATTAATCTCCCTGCAAGTATTCGGTAGTGGATTGAG 2640  
DB 2581 AAGCCCAAGGCTTTAGTAAACCAATGATTAATCTCCCTGCAAGTATTCGGTAGTGGATTGAG 2640  
QY 2641 GAGGACACAAACCCAGTGGAGGCTTTAGTGAAGATCTCAGAGATTTCAATGG 2700  
DB 2641 GAGGACACAAACCCAGTGGAGGCTTTAGTGAAGATCTCAGAGATTTCAATGG 2700  
QY 2701 AGGCGCTGTCTTTTATATACCCAGCTGTACTAGCCCTTATCTGTGMYTTCCCAATATC 2760  
DB 2701 AGGCGCTGTCTTTTATATACCCAGCTGTACTAGCCCTTATCTGTGMYTTCCCAATATC 2760  
QY 2761 CAGAGGAAGCAGAGTGGTTTACASTCTGGAACCTTMAAGATGCTTCTTCTGATCCCTG 2820  
DB 2761 CAGAGGAAGCAGAGTGGTTTACASTCTGGAACCTTMAAGATGCTTCTTCTGATCCCTG 2820  
QY 2821 TACATCTCACTCTCAATTTCTTTGTTGGCTTTGAAGATACTTCAAAACCCARCATCTCAAC 2880  
DB 2821 TACATCTCACTCTCAATTTCTTTGTTGGCTTTGAAGATACTTCAAAACCCARCATCTCAAC 2880  
QY 2881 TCACCTGGAATTTTACCCCAAGGGTTTCAAGGATAGYCCCCATCTATTTGGCCAGGCAT 2940  
DB 2881 TCACCTGGAATTTTACCCCAAGGGTTTCAAGGATAGYCCCCATCTATTTGGCCAGGCAT 2940  
QY 2941 TAGCCCCAGACTTGAGYCARITYMTATACCTGGAACCTTGTCTTCTTCTGATGCTGATG 3000  
DB 2941 TAGCCCCAGACTTGAGYCARITYMTATACCTGGAACCTTGTCTTCTTCTGATGCTGATG 3000  
QY 3001 ATTTACTTTTTRGCGCYRVTTCAGAAAACCTTGTGCACTCAAGCCACCCAGGCTCTTMA 3060  
DB 3001 ATTTACTTTTTRGCGCYRVTTCAGAAAACCTTGTGCACTCAAGCCACCCAGGCTCTTMA 3060  
QY 3061 ATTTCTCTGCACTGTGGCTACAGGTTTCCAAACACACARCTCTGCTCACAGC 3120  
DB 3061 ATTTCTCTGCACTGTGGCTACAGGTTTCCAAACACACARCTCTGCTCACAGC 3120  
QY 3121 AGGTTAAATACCTTAGGCTTAARATTTATCCAAAGGACCACCGGCTCTAGTGAGGAAYRA 3180

|||||  
3121 AGGTTAAATCTTAGGCTTAARATTTATCCAAAGGCAACARGCCCTCAGTGGAGNAYYA 3180  
QY TCCAGCCTATCTGGCTTATCCTCATCYCAAAACCCCTAAAGCAACTAAGRGRRTTCCTTG 3240  
Db TCCAGCCTATCTGGCTTATCCTCATCYCAAAACCCCTAAAGCAACTAAGRGRRTTCCTTG 3240  
QY GCRTAAYAGGYTCTGCCGAAWATGGATTCGCCAGGTWGGCRAATAGCCAGGYCATTA 3300  
Db GCRTAAYAGGYTCTGCCGAAWATGGATTCGCCAGGTWGGCRAATAGCCAGGYCATTA 3300  
QY WATACASTAAATTAAGGAACTCAGAAAGCCTAATACCCATTTARTAAAGATGGAYAMCTGAA 3360  
Db WATACASTAAATTAAGGAACTCAGAAAGCCTAATACCCATTTARTAAAGATGGAYAMCTGAA 3360  
QY GYMAAGTGGCTTCCAGGCCCTTAAAGAGGCCCTTAAACCCCAAGYCCAGGTGTTAAGYT 3420  
Db GYMAAGTGGCTTCCAGGCCCTTAAAGAGGCCCTTAAACCCCAAGYCCAGGTGTTAAGYT 3420  
QY TGCCAAACRGCGCAAGACTTTTSTTATATRTACAGAAAGAAACAGRAAYAGCTCTRGA 3480  
Db TGCCAAACRGCGCAAGACTTTTSTTATATRTACAGAAAGAAACAGRAAYAGCTCTRGA 3480  
QY GTCCTTTACACAGRTCCRAGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYT 3540  
Db GTCCTTTACACAGRTCCRAGGAYGAGCTTGCAACCYRTGGCRYACCTGASTAAGGAAAYT 3540  
QY GATGTAGTGGCAAGGGTTGRCYTCAATTTGTTAYGGGTAGTGGTGGCAGTACAGTYKTA 3600  
Db GATGTAGTGGCAAGGGTTGRCYTCAATTTGTTAYGGGTAGTGGTGGCAGTACAGTYKTA 3600  
QY GTATCTGAAGCAGTTAAAAATAACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
Db GTATCTGAAGCAGTTAAAAATAACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
QY GTGAAYRGCTACTCACTGTCTAAAGGAGACTTGTGCTGTCAAGCAACVGTTTACTTAAA 3720  
Db GTGAAYRGCTACTCACTGTCTAAAGGAGACTTGTGCTGTCAAGCAACVGTTTACTTAAA 3720  
QY TTTCTAGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACTTTGTGCAACTCTTAAC 3780  
Db TTTCTAGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACTTTGTGCAACTCTTAAC 3780  
QY CCAGTNCATTTCTTCCAGA CAATGAAGAAAGATARAAYATACTGTGCAACAARTATTT 3840  
Db CCAGTNCATTTCTTCCAGA CAATGAAGAAAGATARAAYATACTGTGCAACAARTATTT 3840  
QY TCTCAAACTATGCCACTCGAGGGACCTTGTAGARGTTCCVTTGACTGATCCYACCTT 3900  
Db TCTCAAACTATGCCACTCGAGGGACCTTGTAGARGTTCCVTTGACTGATCCYACCTT 3900  
QY CAACCTTGATATCTGATGGAAGTTCCCTTTGTAGAAAAGGACTTCGAAAAGYGGGGTATGC 3960  
Db CAACCTTGATATCTGATGGAAGTTCCCTTTGTAGAAAAGGACTTCGAAAAGYGGGGTATGC 3960  
QY AGTGTGCTGATGATGGAATAYTTGAAAGTAAATCCCTCTCACTCCAGNACTAGTGTCTYA 4020  
Db AGTGTGCTGATGATGGAATAYTTGAAAGTAAATCCCTCTCACTCCAGNACTAGTGTCTYA 4020  
QY GCTRCAGAACTAATAGCCVTCATYTKGGCACTAGAAATTTAGGAGAAAGRAAAAAGGYAAA 4080  
Db GCTRCAGAACTAATAGCCVTCATYTKGGCACTAGAAATTTAGGAGAAAGRAAAAAGGYAAA 4080  
QY TATATATACAGACTCTARTATGCTYACCTAGTTCNTCCATGCCCATGMRGCAATATGSAR 4140  
Db TATATATACAGACTCTARTATGCTYACCTAGTTCNTCCATGCCCATGMRGCAATATGSAR 4140  
QY AGAAAGGGAATTCCTTAACCTTCGAGRGAAACA CTTATCAACATCAGGAAGCCATTAGGAR 4200  
Db AGAAAGGGAATTCCTTAACCTTCGAGRGAAACA CTTATCAACATCAGGAAGCCATTAGGAR 4200  
QY ATTATTATYTGCGWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCGGGGTATCA 4260

Db ATTATTATYTGCGWGTACAGAAACCTARAGAGGTGGMAGTCTTACACTGCGGGGTATCA 4260  
QY NAAAGGAAAGBAAGGGAAATASAAHGAAYTSCCAAGCAKATATTTGAAGCMAAAGAGC 4320  
Db NAAAGGAAAGBAAGGGAAATASAAHGAAYTSCCAAGCAKATATTTGAAGCMAAAGAGC 4320  
QY TGCAAGGCAGACCCCTCCATTAGAAATCTTATTAAACTTCCCTTAGTATATAGGGTAATCC 4380  
Db TGCAAGGCAGACCCCTCCATTAGAAATCTTATTAAACTTCCCTTAGTATATAGGGTAATCC 4380  
QY CTTCCGGGAAACCAAGGCCCTAGTACTCAGCAGGAGAAA CAGAAATGGGAAACCTCAGAGG 4440  
Db CTTCCGGGAAACCAAGGCCCTAGTACTCAGCAGGAGAAA CAGAAATGGGAAACCTCAGAGG 4440  
QY CAGTTTTCTCCCTCGGGACGGTTAGCCACTGAAAGAGGGAANAATCTTTTGCCTGCAAC 4500  
Db CAGTTTTCTCCCTCGGGACGGTTAGCCACTGAAAGAGGGAANAATCTTTTGCCTGCAAC 4500  
QY TATCCAAATGGAAATTA CTTTAAACCCCTTCATCAAAACCTTTTCACTTAGGCATCGATAGCAC 4560  
Db TATCCAAATGGAAATTA CTTTAAACCCCTTCATCAAAACCTTTTCACTTAGGCATCGATAGCAC 4560  
QY CCATCARATGGCCAAATCATTATTTACTGACCAAGGCCCTTTTCAAAACCTATCAAGCANAT 4620  
Db CCATCARATGGCCAAATCATTATTTACTGACCAAGGCCCTTTTCAAAACCTATCAAGCANAT 4620  
QY AKTCAGGCCCTGTGAATGTGCCARABAAAATAATCCCTGCTCTATCCCAAGCTCCTTC 4680  
Db AKTCAGGCCCTGTGAATGTGCCARABAAAATAATCCCTGCTCTATCCCAAGCTCCTTC 4680  
QY AGGARAACAAARAACAGGCCATTCCCTGRRARAARACTTGGCAACTGATTTTACCCACAAG 4740  
Db AGGARAACAAARAACAGGCCATTCCCTGRRARAARACTTGGCAACTGATTTTACCCACAAG 4740  
QY CCCAAACCTCAGGGATTTTCTAGTCTGGGTARATACATTTTCA CGGGTTGGCA 4800  
Db CCCAAACCTCAGGGATTTTCTAGTCTGGGTARATACATTTTCA CGGGTTGGCA 4800  
QY RAGGCTTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACCTAGTTCATGAA 4860  
Db RAGGCTTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACCTAGTTCATGAA 4860  
QY ATAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGA CAATAGCCCTCTTCCAG 4920  
Db ATAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGA CAATAGCCCTCTTCCAG 4920  
QY GCCACAGTAAACCCAGGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCC 4980  
Db GCCACAGTAAACCCAGGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCC 4980  
QY TGAAGGCCACAGTCTCTCAGGGAAGTTCGAGAAAATGAATGAAYACTCAAGGACATCTA 5040  
Db TGAAGGCCACAGTCTCTCAGGGAAGTTCGAGAAAATGAATGAAYACTCAAGGACATCTA 5040  
QY AAAAGCAAAACCCAGGAAACCCACCTCAGTGGCTGTCTGTGCTATAGCCTTAAAA 5100  
Db AAAAGCAAAACCCAGGAAACCCACCTCAGTGGCTGTCTGTGCTATAGCCTTAAAA 5100  
QY AGAATCTGCAACTTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAG 5160  
Db AGAATCTGCAACTTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAG 5160  
QY CCCTTCATAACCAATGACTTTGTGCTTGACCCCAAGACCAACTTAGTTCAGACATCA 5220  
Db CCCTTCATAACCAATGACTTTGTGCTTGACCCCAAGACCAACTTAGTTCAGACATCA 5220  
QY CCTCTCTAGCCAAATATCAACAGTTCCTTAAACAATTACAGAAACCTATCCCTGAGAAG 5280  
Db CCTCTCTAGCCAAATATCAACAGTTCCTTAAACAATTACAGAAACCTATCCCTGAGAAG 5280  
QY AGGAAAAAGAACTATTTCACCCCMWGTGACATGGTATAGTCAAGTCCCTTCCTCTAAAT 5340  
Db AGGAAAAAGAACTATTTCACCCCMWGTGACATGGTATAGTCAAGTCCCTTCCTCTAAAT 5340

QY	5341	CCCATCCCTAGATACATCTCTGGGAGGACCCCTACCCAGTCATTTATTTATCCCAACTG	5400
Db	5341		
QY	5401	CGGTTAAAGTGGCTGGAGTGGAGTCTTTGGATACATCACTTGGAGCAAACTCTGGATAC	5460
Db	5401	CGGTTAAAGTGGCTGGAGTGGAGTCTTTGGATACATCACTTGGAGCAAACTCTGGATAC	5460
QY	5461	TGCCAAGGAACTGGAATCCAGGAGACAACCGTAGCTATTTCTGTGAACCTCTAGAGG	5520
Db	5461	TGCCAAGGAACTGGAATCCAGGAGACAACCGTAGCTATTTCTGTGAACCTCTAGAGG	5520
QY	5521	ATTTGGCCCTGCTCTTCAAAACAACACGAGGAGAAAGTAACATAAATCATAAATCCCC	5580
Db	5521	ATTTGGCCCTGCTCTTCAAAACAACACGAGGAGAAAGTAACATAAATCATAAATCCCC	5580
QY	5581	ATGSGCTCCCTTATCATATTTTCTCTASTGTSTTTTACCCSTTTTCACTCTCACT	5640
Db	5581	ATGSGCTCCCTTATCATATTTTCTCTASTGTSTTTTACCCSTTTTCACTCTCACT	5640
QY	5641	GCACCCCTCCATGCGCTGTATGACGAGTAGCTCCCTYACMWAGAGTTTCTATGGAGA	5700
Db	5641	GCACCCCTCCATGCGCTGTATGACGAGTAGCTCCCTYACMWAGAGTTTCTATGGAGA	5700
QY	5701	ATGCAGCTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGAAACCC	5760
Db	5701	ATGCAGCTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGAAACCC	5760
QY	5761	ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG	5820
Db	5761	ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG	5820
QY	5821	CATGCAAACTCATTTATGGACAGGAAATGATTAATCTAGTTGTCTGGAGGACTT	5880
Db	5821	CATGCAAACTCATTTATGGACAGGAAATGATTAATCTAGTTGTCTGGAGGACTT	5880
QY	5881	GGAGTCACTGTCTGTGGACTTACTTCAACCAACTGGTATGTCGATGGGGTGAGTT	5940
Db	5881	GGAGTCACTGTCTGTGGACTTACTTCAACCAACTGGTATGTCGATGGGGTGAGTT	5940
QY	5941	CAAGATCAGCAAGAGAAACATGTAAAGAGTAATCTCCCAACTCAACCCGGGTACAT	6000
Db	5941	CAAGATCAGCAAGAGAAACATGTAAAGAGTAATCTCCCAACTCAACCCGGGTACAT	6000
QY	6001	GGACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCCTCCGTACC	6060
Db	6001	GGACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCCTCCGTACC	6060
QY	6061	CATACCTCGCTGGTAGCCCTATTTAATACACCTCACTGGGCTCCATGAGGTCTCGGCC	6120
Db	6061	CATACCTCGCTGGTAGCCCTATTTAATACACCTCACTGGGCTCCATGAGGTCTCGGCC	6120
QY	6121	CAAAACCTTAACTGTTGGATATGCTCCCTCGAACTTCARGCCATATGTTTCAATC	6180
Db	6121	CAAAACCTTAACTGTTGGATATGCTCCCTCGAACTTCARGCCATATGTTTCAATC	6180
QY	6181	CTGTACCTGAAACAATGGAACTTTCAGCACAGAAATAAACACCACTTCCGTTTATGTA	6240
Db	6181	CTGTACCTGAAACAATGGAACTTTCAGCACAGAAATAAACACCACTTCCGTTTATGTA	6240
QY	6241	GGACCTCTGTTTCCAAATSTGGAATTAACCATACCTCAAACTCACTGTGTGAAATTTT	6300
Db	6241	GGACCTCTGTTTCCAAATSTGGAATTAACCATACCTCAAACTCACTGTGTGAAATTTT	6300
QY	6301	AGCAATCTACATACACCAACTCCCAATGCATGAGTGGGTAACTCTCCCAACAA	6360
Db	6301	AGCAATCTACATACACCAACTCCCAATGCATGAGTGGGTAACTCTCCCAACAA	6360
QY	6361	ATAGTCTGCTACCTCAGAAATATTTTGTCTGTGGTACCTCAGCCCTATCGTTGTTTG	6420
Db	6361	ATAGTCTGCTACCTCAGAAATATTTTGTCTGTGGTACCTCAGCCCTATCGTTGTTTG	6420

QY	6421	AATGGCTCTTCAGAAATCTATGTGCTTCTCTCATTTCTAGTGCCCCCYATGRCATCTAC	6480
Db	6421	AATGGCTCTTCAGAAATCTATGTGCTTCTCTCATTTCTAGTGCCCCCYATGRCATCTAC	6480
QY	6481	ACTGAAACAAGATTTATACAGTTTATGTCAATCTAAAGCCCCGCAACAAAGAGTACCCATT	6540
Db	6481	ACTGAAACAAGATTTATACAGTTTATGTCAATCTAAAGCCCCGCAACAAAGAGTACCCATT	6540
QY	6541	CTTCTCTTTTGTATAGGAGGAGGTGCTAGTGCACTAGGTACTGGCATTTGGCGGTATC	6600
Db	6541	CTTCTCTTTTGTATAGGAGGAGGTGCTAGTGCACTAGGTACTGGCATTTGGCGGTATC	6600
QY	6601	ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGACCG	6660
Db	6601	ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGACCG	6660
QY	6661	GTGCGCGACTCCCTGCTCACCTTTCGAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTT	6720
Db	6661	GTGCGCGACTCCCTGCTCACCTTTCGAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTT	6720
QY	6721	CRAAATCGAAGAGCTTTTAGACTTTCGCTAACCGCTGABAGAGGGGAACTGTTTATTTT	6780
Db	6721	CRAAATCGAAGAGCTTTTAGACTTTCGCTAACCGCTGABAGAGGGGAACTGTTTATTTT	6780
QY	6781	GGGGAAGATGCTGTTTATTTATTTATCAATCGGAACTCGTCACTGAGAAAGTTAAGAA	6840
Db	6781	GGGGAAGATGCTGTTTATTTATTTATCAATCGGAACTCGTCACTGAGAAAGTTAAGAA	6840
QY	6841	ATTCAGATCGAATCAAGCTAKAGCAGARGAGCTTCGAAACACTCGACCTCGGGCCTC	6900
Db	6841	ATTCAGATCGAATCAAGCTAKAGCAGARGAGCTTCGAAACACTCGACCTCGGGCCTC	6900
QY	6901	CTCAGCCRAATGATGCCCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATTG	6960
Db	6901	CTCAGCCRAATGATGCCCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATTG	6960
QY	6961	CTACTCTCTTTGGACCTCTGATCTTTTACCTCTCTGTTTAACTTTCTCTCCAGATC	7020
Db	6961	CTACTCTCTTTGGACCTCTGATCTTTTACCTCTCTGTTTAACTTTCTCTCCAGATC	7020
QY	7021	GAAGCTGTAAACTCAAAATGAGGCCCAAGATGCACTGAGTCCCAAGACTAGTCCAGAG	7080
Db	7021	GAAGCTGTAAACTCAAAATGAGGCCCAAGATGCACTGAGTCCCAAGACTAGTCCAGAG	7080
QY	7081	CCCCGGGACCGGCTGTAGCCCAAGATCTGATGTTTAACTGCAATCAAGGACCCCTCCT	7140
Db	7081	CCCCGGGACCGGCTGTAGCCCAAGATCTGATGTTTAACTGCAATCAAGGACCCCTCCT	7140
QY	7141	GAGGAAATCTCAGCTGCAAACTCTCTA CTACGCCCCAATTCAGCAGGAGAGGTTAGAGC	7200
Db	7141	GAGGAAATCTCAGCTGCAAACTCTCTA CTACGCCCCAATTCAGCAGGAGAGGTTAGAGC	7200
QY	7201	GGTGTGGGCAACCTCCCAACAGCAGCTTAGGTTTCTGTTGATGGGGACCTGAGA	7260
Db	7201	GGTGTGGGCAACCTCCCAACAGCAGCTTAGGTTTCTGTTGATGGGGACCTGAGA	7260
QY	7261	GACGAGCTAGCTGGATTTCTTAGCTGATTAAGAACTCCYTAAGCCTAGTGTGGGAGGTG	7320
Db	7261	GACGAGCTAGCTGGATTTCTTAGCTGATTAAGAACTCCYTAAGCCTAGTGTGGGAGGTG	7320
QY	7321	ACCAATCACTTTTAAACAACCGGGCTTGCACTTAGYTCACACCTGACCAATCAGAGAG	7380
Db	7321	ACCAATCACTTTTAAACAACCGGGCTTGCACTTAGYTCACACCTGACCAATCAGAGAG	7380
QY	7381	CTCACTAAATGCTAATTTAGGCAAAAGAGGAGTTAAGAAATAGCCAAATCATYATTGC	7440
Db	7381	CTCACTAAATGCTAATTTAGGCAAAAGAGGAGTTAAGAAATAGCCAAATCATYATTGC	7440
QY	7441	MTGAGGACACACGAGGAGCAATGATCGGGATATAACCAAGTTTTCGAGCCGGCAA	7500
Db	7441	MTGAGGACACACGAGGAGCAATGATCGGGATATAACCAAGTTTTCGAGCCGGCAA	7500
QY	7501	CGGCAACCCCTTTGGTCCCTCTCTGTTGATGGAGCTCTGTTTTCATGCTATTTCAC	7560

Db	7501	CGGCAACCCCTTTGGGTGCCCTCCCTTTGTATGGAGCTCTGTGTTTCATGCTATTTTCAC	7560
Qy	7561	TCATTAAATCTTGCARCTGCR	7582
Db	7561	TCATTAAATCTTGCARCTGCR	7582
RESULT 2			
AAX25665			
ID	AAX25665 standard; cDNA to mRNA; 7582 BP.		
XX	AC	AAX25665;	
XX	XX		
DT	21-MAY-1999 (first entry)		
XX	XX		
DE	Complete human endogenous retrovirus W genome.		
XX	XX		
KW	Clone; human endogenous retrovirus; genome; autoimmune disease;		
KW	multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;		
KW	disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.		
XX	XX		
OS	Human endogenous retrovirus.		
XX	XX		
FN	WO9902696-A1.		
XX	XX		
PD	21-JAN-1999.		
XX	XX		
PF	06--JUL-1998; 98WO-FR001442.		
XX	XX		
PR	07--JUL-1997; 97FR-00008815.		
XX	XX		
PA	(INMR ) BIO MERIEUX.		
XX	XX		
FI	Beseme F, Blond J, Bouton O, Mandrand B, Mallet F;		
XX	XX		
DR	WPI; 1999-120897/10.		
PT	New nucleic acid sequences from human endogenous retrovirus-W - expressed		
PT	exclusively in placenta and useful in diagnosis and therapy of autoimmune		
PT	disease, and abnormal or failed pregnancy.		
XX	XX		
PS	Claim 1; Page 71-74; 106pp; French.		
XX	XX		
CC	This sequence represents the complete sequence of the human endogenous		
CC	retrovirus (HERV) W genome. The nucleic acids, their fragments or		
CC	peptides encoded by them are markers of autoimmune disease (e.g. multiple		
CC	sclerosis, rheumatoid polyarthritis, disseminated lupus erythematosus,		
CC	insulin- dependent diabetes and related pathologies) and of abnormal or		
CC	unsuccessful pregnancy and can be used as chromosomal markers for		
CC	susceptibility to these conditions, or proximity markers of genes		
CC	associated with this susceptibility		
XX	XX		
SQ	Sequence 7582 BP; 2156 A; 1877 C; 1537 G; 1796 T; 2 U; 214 Other;		
Query Match 97.3%; Score 7378; DB 2; Length 7582;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 7578; Conservative 0; Mismatches 4; Indels 0; Gaps 0;			
Qy	1	CAACAAATCGGGATATAAACCCAGGCAATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCC	60
Db	1	CAACAAATCGGGATATAAACCCAGGCAATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCC	60
Qy	61	CTTCCCTTTGTATGGAGCTGTTTTCAATCTATTTCACTATTTAAATCTTGCAACTGCA	120
Db	61	CTTCCCTTTGTATGGAGCTGTTTTCAATCTATTTCACTATTTAAATCTTGCAACTGCA	120
Qy	121	CTCTTCTGGTCCATGTTTTCTTACGGCTCGAGCTGAGCTTTTGTCTCACGGTCCACCACTGC	180
Db	121	CTCTTCTGGTCCATGTTTTCTTACGGCTCGAGCTGAGCTTTTGTCTCACGGTCCACCACTGC	180
Qy	181	TGTTTGCCACCAACCGCANACTGGCGGTGACTCCCATCCCTCTGGATCCTCAGGGTGC	240



QY 1321 GAAGGAAGANAATGATTTCCCAAGGCCAGCAGCGAGTCCAGTCTASACCTCATTTG 1380  
DB 1321 GAAGGAAGANAATGATTTCCCAAGGCCAGCAGCGAGTCCAGTCTASACCTCATTTG 1380  
QY 1381 GGGACACAGAAATCAGTAACATGGGAGATTTGGTGTGTCAGACATTTTCTAACTTTGTGTGC 1440  
DB 1381 GGGACACAGAAATCAGTAACATGGGAGATTTGGTGTGTCAGACATTTTCTAACTTTGTGTGC 1440  
QY 1441 TASAAGGACTAAGGAATACTTASGAAGAAATCTTAYGAATTTACTCAATGATGTCCACCATA 1500  
DB 1441 TASAAGGACTAAGGAATACTTASGAAGAAATCTTAYGAATTTACTCAATGATGTCCACCATA 1500  
QY 1501 ACACAGGGGAGGAGAGAAATCTTACTGCTTTCTGGAGAGACTAAGGAGGCAATTGAG 1560  
DB 1501 ACACAGGGGAGGAGAGAAATCTTACTGCTTTCTGGAGAGACTAAGGAGGCAATTGAG 1560  
QY 1561 GAAGCGTGCCTCTCTGTCACTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAAGTT 1620  
DB 1561 GAAGCGTGCCTCTCTGTCACTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAAGTT 1620  
QY 1621 TATCACTCAGTGCAGACATTAAGAAAACTTCAAAAGTCTGCGGTAGGCCCGGAG 1680  
DB 1621 TATCACTCAGTGCAGACATTAAGAAAACTTCAAAAGTCTGCGGTAGGCCCGGAG 1680  
QY 1681 CAAAACCTTAGAACCCTATTGAATTTGGCAACCTCGGTTTTTATATAGAGATCAGGAG 1740  
DB 1681 CAAAACCTTAGAACCCTATTGAATTTGGCAACCTCGGTTTTTATATAGAGATCAGGAG 1740  
QY 1741 GAGCAGCGGAAACAGGACAAACGGGATTAAGAAAAAGGCCACCGCTTTAGTCATGACCCT 1800  
DB 1741 GAGCAGCGGAAACAGGACAAACGGGATTAAGAAAAAGGCCACCGCTTTAGTCATGACCCT 1800  
QY 1801 CAGGCAAGTGGACTTTGGAGGCTCTGGAAGGGAAGAAAGCTGGGCAATTTGAATGCTTAA 1860  
DB 1801 CAGGCAAGTGGACTTTGGAGGCTCTGGAAGGGAAGAAAGCTGGGCAATTTGAATGCTTAA 1860  
QY 1861 TAGGGCTGCTTCCAGTGGGTCTACAAGACACTTTAAAGAGATGTCCAAAGTAGAAG 1920  
DB 1861 TAGGGCTGCTTCCAGTGGGTCTACAAGACACTTTAAAGAGATGTCCAAAGTAGAAG 1920  
QY 1921 TAAGCGCCCTCTGCTCCATGCCCTTTATTTCAAGGGAATCACTGGAAGGCCACCTGCCC 1980  
DB 1921 TAAGCGCCCTCTGCTCCATGCCCTTTATTTCAAGGGAATCACTGGAAGGCCACCTGCCC 1980  
QY 1981 CAGGGACAAAGGTCTTTTGAAGTCAAGACCACTAACAGATGATCCAGCAGCAGACTG 2040  
DB 1981 CAGGGACAAAGGTCTTTTGAAGTCAAGACCACTAACAGATGATCCAGCAGCAGACTG 2040  
QY 2041 AGGGTGCCTGGGCAAGCGCATCCCATGCCATCACTCACAGAGCCCTGGGTATGCTTT 2100  
DB 2041 AGGGTGCCTGGGCAAGCGCATCCCATGCCATCACTCACAGAGCCCTGGGTATGCTTT 2100  
QY 2101 GACCATTTGAGGGCCAGGAGGTGTCTCTGACACTGGTGGGTCTTCTAGTCTTACT 2160  
DB 2101 GACCATTTGAGGGCCAGGAGGTGTCTCTGACACTGGTGGGTCTTCTAGTCTTACT 2160  
QY 2161 CTTCCTGCTCCGGACAACTGTCTCCAGATCTGTCACTATTTCTGAGGGGTCCNTAAGAG 2220  
DB 2161 CTTCCTGCTCCGGACAACTGTCTCCAGATCTGTCACTATTTCTGAGGGGTCCNTAAGAG 2220  
QY 2221 GGCAGTCACTAGATACCTTTTCCAGCCACTAAGTTTATGAACCTGGGGAGCTTTATTTCTTT 2280  
DB 2221 GGCAGTCACTAGATACCTTTTCCAGCCACTAAGTTTATGAACCTGGGGAGCTTTATTTCTTT 2280  
QY 2281 TCACATGCTTTCTAATATGCTTTGAAGCCCACTACCTTGTAGGGAGAGACATTTCTA 2340  
DB 2281 TCACATGCTTTCTAATATGCTTTGAAGCCCACTACCTTGTAGGGAGAGACATTTCTA 2340  
QY 2341 GCAAAAGCAGGGCCATTAACACCTGAACATAGGAGAGGAGAACCCCGTTTGTCTGCC 2400  
DB 2341 GCAAAAGCAGGGCCATTAACACCTGAACATAGGAGAGGAGAACCCCGTTTGTCTGCC 2400

QY 2401 CCTGCTTGAGGAAGAAATTAATCCTCAAGTCTGGGCAACAGAGGACAAATATGACGAGC 2460  
DB 2401 CCTGCTTGAGGAAGAAATTAATCCTCAAGTCTGGGCAACAGAGGACAAATATGACGAGC 2460  
QY 2461 CAAAGAAATGCCGCTCTGTTCAAGTTAACTAAGGATTTCCACTTCTTCCCTTCCCTACCAA 2520  
DB 2461 CAAAGAAATGCCGCTCTGTTCAAGTTAACTAAGGATTTCCACTTCTTCCCTTCCCTACCAA 2520  
QY 2521 GGCAGTACCCCTCAGACCCCAAGGCCCAACAGGATTTCCAAAAGATTTGTTAAGGACTTAA 2580  
DB 2521 GGCAGTACCCCTCAGACCCCAAGGCCCAACAGGATTTCCAAAAGATTTGTTAAGGACTTAA 2580  
QY 2581 AAGCCCAAGGCTTAGTAAACCATGTATACCTCCCTGCGAGTAATTCCTGTAGTGTAG 2640  
DB 2581 AAGCCCAAGGCTTAGTAAACCATGTATACCTCCCTGCGAGTAATTCCTGTAGTGTAG 2640  
QY 2641 GAGGACACAGAAACCCAGTGGAGGAGTGTAGTGCAGATCTCAGGATTAATCAATGG 2700  
DB 2641 GAGGACACAGAAACCCAGTGGAGGAGTGTAGTGCAGATCTCAGGATTAATCAATGG 2700  
QY 2701 AGGCCGTTGCTCTTTTATACCCAGCTGTACTAGCCCTTATATCTGTGMYTTTCCCAATAC 2760  
DB 2701 AGGCCGTTGCTCTTTTATACCCAGCTGTACTAGCCCTTATATCTGTGMYTTTCCCAATAC 2760  
QY 2761 CAGAGGAAGGAGTGGTTTACASTCCTGAGCCTTMMAGGATGCCCTTCTCTGATCCCTG 2820  
DB 2761 CAGAGGAAGGAGTGGTTTACASTCCTGAGCCTTMMAGGATGCCCTTCTCTGATCCCTG 2820  
QY 2821 TACATCCTGACTCAATTTCTGTTTGGCTTTGAAGATTAATTTCAAAACCCCAATCTCAAC 2880  
DB 2821 TACATCCTGACTCAATTTCTGTTTGGCTTTGAAGATTAATTTCAAAACCCCAATCTCAAC 2880  
QY 2881 TCACCTGGAATTTTACCCCAAGGTTTCAAGGATAGYCCCATCTATTTGGGCGAGCAT 2940  
DB 2881 TCACCTGGAATTTTACCCCAAGGTTTCAAGGATAGYCCCATCTATTTGGGCGAGCAT 2940  
QY 2941 TAGCCCAAGACTTGAGYCATYMTCATAGCTGGACACTTGTCTTCTGATGTGATG 3000  
DB 2941 TAGCCCAAGACTTGAGYCATYMTCATAGCTGGACACTTGTCTTCTGATGTGATG 3000  
QY 3001 ATTTACTTTTTCGCGCCYCTTTCAGAAACCTTGTGCCATCAAGCCACCCAGGCTCTTMA 3060  
DB 3001 ATTTACTTTTTCGCGCCYCTTTCAGAAACCTTGTGCCATCAAGCCACCCAGGCTCTTMA 3060  
QY 3061 ATTTCTCTCCYACCTTGTGGCTACAWGGTTTCCAAACSAARGCTCACTCTGCTCAGC 3120  
DB 3061 ATTTCTCTCCYACCTTGTGGCTACAWGGTTTCCAAACSAARGCTCACTCTGCTCAGC 3120  
QY 3121 AGGTTAAATATTTAGGRTTAABATTATCCAAAGGCCACCCARGGCCCTCAGTGAAGAA 3180  
DB 3121 AGGTTAAATATTTAGGRTTAABATTATCCAAAGGCCACCCARGGCCCTCAGTGAAGAA 3180  
QY 3181 TCCAGGCTTATCTGGCTTATCTCATCYCAAAACCTTAAGCAACTAAGGRRTTCTTGT 3240  
DB 3181 TCCAGGCTTATCTGGCTTATCTCATCYCAAAACCTTAAGCAACTAAGGRRTTCTTGT 3240  
QY 3241 GCRTAAYAGGTTCTGCGAAWATGATTTCCCGAGTWTGGCRAAATAGCCAGGYCATTA 3300  
DB 3241 GCRTAAYAGGTTCTGCGAAWATGATTTCCCGAGTWTGGCRAAATAGCCAGGYCATTA 3300  
QY 3301 WATACASTAATTAAGGAACTCAGAAAGCCAAATACCATTTTARTAGATGGAYAMCTGAA 3360  
DB 3301 WATACASTAATTAAGGAACTCAGAAAGCCAAATACCATTTTARTAGATGGAYAMCTGAA 3360  
QY 3361 GWMRAAGTGGCTTTCAGGCCCTTAAAGAGGCCCTTAAACCCAAAGYCCCAAGTGTAA 3420  
DB 3361 GWMRAAGTGGCTTTCAGGCCCTTAAAGAGGCCCTTAAACCCAAAGYCCCAAGTGTAA 3420  
QY 3421 TGCCCAACGGGCAAGACTTTTSTTATATYRTACAGAAAGAAACAGRAAYAGCTCTRGA 3480  
DB 3421 TGCCCAACGGGCAAGACTTTTSTTATATYRTACAGAAAGAAACAGRAAYAGCTCTRGA 3480  
QY 3481 GTCTTTACACAGRTCCRAGGGAYGAGCTTGCACCCVRTGGCRYACCTGTAGGAAAVT 3540

|||||  
3481 GTCTTACACAGTTCRAGGGAAGCTTGCACCCYRTGGCRYACTGASTAAGGAAT 3540  
QY  
3541 GATGTAGTGGCAAGGGTTGRCYTCATTTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTA 3600  
Db  
3541 GATGTAGTGGCAAGGGTTGRCYTCATTTGTTTAYGGGTAGTGGTGGCAGTAGCAGTYKTA 3600  
QY  
3601 GTATCTGAAGCAGTTAAATAATATACAGGGRAGAGATCTTACGTGTGGACATCTCATGAK 3660  
Db  
3601 GTATCTGAAGCAGTTAAATAATATACAGGGRAGAGATCTTACGTGTGGACATCTCATGAK 3660  
QY  
3661 GTGAAYRGCACTACTCACTCTAAAGSAGACTTGTGCTGTACAGACAACVGTTTACTTAAA 3720  
Db  
3661 GTGAAYRGCACTACTCACTCTAAAGSAGACTTGTGCTGTACAGACAACVGTTTACTTAAA 3720  
QY  
3721 TRTCAGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACTTTGTGCAACTCTTAAC 3780  
Db  
3721 TRTCAGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACTTTGTGCAACTCTTAAC 3780  
QY  
3781 CCAGYCNCAATTTCTCCAGACAATGAAGAAAGATARAAYATAA CTGTCTCAACAATTAAT 3840  
Db  
3781 CCAGYCNCAATTTCTCCAGACAATGAAGAAAGATARAAYATAA CTGTCTCAACAATTAAT 3840  
QY  
3841 TCTCAAAACCTATGCCACTCGAGGGACCTTGTAGAGTTCCYTTGACTGATCCYACCTT 3900  
Db  
3841 TCTCAAAACCTATGCCACTCGAGGGACCTTGTAGAGTTCCYTTGACTGATCCYACCTT 3900  
QY  
3901 CAACTTGTATCTAGTGAAGTTCCCTTGTAGAAAAGGACTTCGAAAAGYGGGGTATGC 3960  
Db  
3901 CAACTTGTATCTAGTGAAGTTCCCTTGTAGAAAAGGACTTCGAAAAGYGGGGTATGC 3960  
QY  
3961 AGTGGTCAGTGAATAATGGAATATTTGAAAGTAATCCCTCACTCAGGAACTAGTGCTYA 4020  
Db  
3961 AGTGGTCAGTGAATAATGGAATATTTGAAAGTAATCCCTCACTCAGGAACTAGTGCTYA 4020  
QY  
4021 GCTRCGACAACTAATAGCCCTCACTYKGGCACTAGAAATAGGAGAGRAAAGGGYAAA 4080  
Db  
4021 GCTRCGACAACTAATAGCCCTCACTYKGGCACTAGAAATAGGAGAGRAAAGGGYAAA 4080  
QY  
4081 TATATATACAGACTCTTARTATGCTYACCTAGTCTNTCCATGCCATGMRGCAATAGSAR 4140  
Db  
4081 TATATATACAGACTCTTARTATGCTYACCTAGTCTNTCCATGCCATGMRGCAATAGSAR 4140  
QY  
4141 AGAAGGGAATTCCTAACTTCYAGRGAACTATATCAACATCAGGAAGCCATTAGGAR 4200  
Db  
4141 AGAAGGGAATTCCTAACTTCYAGRGAACTATATCAACATCAGGAAGCCATTAGGAR 4200  
QY  
4201 ATTATATYTGCGWGTACAGAACTTARAGAGGTGGWAGTCTTACACTGCGYGGGTATCA 4260  
Db  
4201 ATTATATYTGCGWGTACAGAACTTARAGAGGTGGWAGTCTTACACTGCGYGGGTATCA 4260  
QY  
4261 NAAAGGAAAGRAAAGGGAATASAGRGAAATGCGCAAGCAKATATTTGAAGCMAAAGAGC 4320  
Db  
4261 NAAAGGAAAGRAAAGGGAATASAGRGAAATGCGCAAGCAKATATTTGAAGCMAAAGAGC 4320  
QY  
4321 TGCAAGGCAAGGACCTTCAATAGAAATGCTTATTAACCTTCCCTTAGTATAGGTAATCC 4380  
Db  
4321 TGCAAGGCAAGGACCTTCAATAGAAATGCTTATTAACCTTCCCTTAGTATAGGTAATCC 4380  
QY  
4381 CTTCCGGGAAACCAAGCCCACTACTCAGCAGGAGAAACAGAAATGGGAACTCTCAGGAG 4440  
Db  
4381 CTTCCGGGAAACCAAGCCCACTACTCAGCAGGAGAAACAGAAATGGGAACTCTCAGGAG 4440  
QY  
4441 CAGTTTTCTCCCTCGGAGCGGTAGCCACTGAAGAGGGAATACTTTTGCCTGCAAC 4500  
Db  
4441 CAGTTTTCTCCCTCGGAGCGGTAGCCACTGAAGAGGGAATACTTTTGCCTGCAAC 4500  
QY  
4501 TATCCAATGGAAATTAATTAACCTTCAATCAACCTTTCACTTAGGCATCGATAGCAC 4560  
Db  
4501 TATCCAATGGAAATTAATTAACCTTCAATCAACCTTTCACTTAGGCATCGATAGCAC 4560  
QY  
4561 CCATCATATGGCCAAATCATTTTACTGACACAGGCCCTTTTCAAAACTATCAAGCARAT 4620  
|||||

Db  
4561 CCATCATATGGCCAAATCATTTTACTGACACAGGCCCTTTTCAAACTATCAAGCARAT 4620  
QY  
4621 AKTCHAGGCCCTGTGAATGTGCCAPARAAAATAATCCCTGCCCTYATCGCAGAGCTCCTTC 4680  
Db  
4621 AKTCHAGGCCCTGTGAATGTGCCAPARAAAATAATCCCTGCCCTYATCGCAGAGCTCCTTC 4680  
QY  
4681 AGGABAAACAAAARACAGGCCATTACCTGRRARAARACTGGCAACTGATTTTATCCCAACAG 4740  
Db  
4681 AGGABAAACAAAARACAGGCCATTACCTGRRARAARACTGGCAACTGATTTTATCCCAACAG 4740  
QY  
4741 CCCAAACCTCAGGGAATTTTCACTATCTACTAGTCTGGGTARATACCTTTTCACTGGGTGGGCA 4800  
Db  
4741 CCCAAACCTCAGGGAATTTTCACTATCTACTAGTCTGGGTARATACCTTTTCACTGGGTGGGCA 4800  
QY  
4801 RAGGCTTTCCCTGTAGGACGAAAGGCCCAAGAGGTAAATAAAGGCACATAGTTTCATGAA 4860  
Db  
4801 RAGGCTTTCCCTGTAGGACGAAAGGCCCAAGAGGTAAATAAAGGCACATAGTTTCATGAA 4860  
QY  
4861 ATATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAG 4920  
Db  
4861 ATATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAG 4920  
QY  
4921 GCCCAGTAACCCAGGAGTATCCCAGGCTTGTAGGTATACGATATCACTTACACTGCGCC 4980  
Db  
4921 GCCCAGTAACCCAGGAGTATCCCAGGCTTGTAGGTATACGATATCACTTACACTGCGCC 4980  
QY  
4981 TGAAGGCCACAGTCTCTAGGGAAGTTCAGAAAAATGAATGAAYACTCAAGAGCATCTA 5040  
Db  
4981 TGAAGGCCACAGTCTCTAGGGAAGTTCAGAAAAATGAATGAAYACTCAAGAGCATCTA 5040  
QY  
5041 AAAAAGCAAAACCCAGGAAACCCACCTCACATGCGCTGTCTGTGCTGCTATAGCCCTTAAA 5100  
Db  
5041 AAAAAGCAAAACCCAGGAAACCCACCTCACATGCGCTGTCTGTGCTGCTATAGCCCTTAAA 5100  
QY  
5101 AGAATCTGCAACTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAG 5160  
Db  
5101 AGAATCTGCAACTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAG 5160  
QY  
5161 CCCTTCTAACAATGACTTGTGTGACCCCAAGACAGCCAACTTAGTTGAGACATCA 5220  
Db  
5161 CCCTTCTAACAATGACTTGTGTGACCCCAAGACAGCCAACTTAGTTGAGACATCA 5220  
QY  
5221 CCTCTTAGCCAAATATCAACAGTTCTTAAACATTAACAGAAACCTTATCCCTGAGAAG 5280  
Db  
5221 CCTCTTAGCCAAATATCAACAGTTCTTAAACATTAACAGAAACCTTATCCCTGAGAAG 5280  
QY  
5281 AGGGAAAAAGAACTATTTCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTTAAT 5340  
Db  
5281 AGGGAAAAAGAACTATTTCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTTAAT 5340  
QY  
5341 CCCCATCTCAGATACATCTCTGGGAAGGACCTTACCCAGTCACTTTTATVTAACCAACTG 5400  
Db  
5341 CCCCATCTCAGATACATCTCTGGGAAGGACCTTACCCAGTCACTTTTATVTAACCAACTG 5400  
QY  
5401 CGGTAAAGTGGCTGAGTGGAGTCTTGGATACATCACTTGGAGTCAATCTTGATAC 5460  
Db  
5401 CGGTAAAGTGGCTGAGTGGAGTCTTGGATACATCACTTGGAGTCAATCTTGATAC 5460  
QY  
5461 TGCCAAAAGGAACCTCAAAATCCAGGAGCAACGCTAGCTATTCTGTGAAACCTCTAGAGG 5520  
Db  
5461 TGCCAAAAGGAACCTCAAAATCCAGGAGCAACGCTAGCTATTCTGTGAAACCTCTAGAGG 5520  
QY  
5521 ATTTCCGGCTGCTCTTCAAAAACAAACCAAGGAGGAAGTAACCTAAATCATAAATCCCC 5580  
Db  
5521 ATTTCCGGCTGCTCTTCAAAAACAAACCAAGGAGGAAGTAACCTAAATCATAAATCCCC 5580  
QY  
5581 ATGSGSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCTSTTTCTCACT 5640  
Db  
5581 ATGSGSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCTSTTTCTCACT 5640  
QY  
5641 GCACCCCTCCATGCGGTGTATGACAGTACCTCCCTYACCMAGAGTTTCTATGGAGA 5700  
Db  
5641 GCACCCCTCCATGCGGTGTATGACAGTACCTCCCTYACCMAGAGTTTCTATGGAGA 5700



XX	11-OCT-2001.
PD	
XX	
XX	30-MAR-2001; 2001WO-US008631.
Pf	
XX	
PR	31-MAR-2000; 2000US-00540217.
PR	23-AUG-2000; 2000US-00649167.
XX	(HYSE-) HYSEQ INC.
PA	
XX	
PI	Dzmanac RT, Liu C, Tang YT;
PS	
DR	WPI; 2001-639362/73.
DR	P-PSDB; ABG04439.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity.
XX	
PS	Claim 1; SEQ ID NO 4430; 103pp; English.
XX	
CC	The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC	sequences. (I) is useful as hybridisation probes, polymerase chain
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC	and in recombinant production of (II). The polynucleotides are also used
CC	in diagnostics as expressed sequence tags for identifying expressed
CC	genes. (I) is useful in gene therapy techniques to restore normal
CC	activity of (II) or to treat disease states involving (II). (II) is
CC	useful for generating antibodies against it, detecting or quantitating a
CC	polypeptide in tissue, as molecular weight markers and as a food
CC	supplement. (II) and its binding partners are useful in medical imaging
CC	of sites expressing (II). (I) and (II) are useful for treating disorders
CC	involving aberrant protein expression or biological activities. The
CC	polypeptide and polynucleotide sequences have applications in:
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC	coding sequences of the invention. Note: The sequence data for this
CC	patent did not appear in the printed specification, but was obtained in
CC	electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 7466 BP; 2217 A; 1880 C; 1614 G; 1754 T; 0 U; 1 Other;
Query Match            12.9%; Score 979; DB 5; Length 7466;	
Best Local Similarity 99.8%; Pred. No. 0;	
Matches 1269; Conservative 0; Mismatches 1; Indels 2; Gaps 2;	
Qy	1476 GAATTACTCAATGATGTCCACCATAACACAGGGGGAAGAAATCCTTACTGCCTTC 1535
Db	
Yq	5407 GAATTACTCAATGATGTCCAATACACAGGGGGAAGAAATCCTTACTGCCTTC 5466
Qy	1536 TGGAGAGACTAAGGGAGCGCATTGAGGAAGCGTGCTCTCTGTCACTGACTCTTCTGAAG 1595
Db	
Yq	5467 TGGNAGACTAAGGGAGCGCATTGAGGAAGCGTGCTCTCTGTCACTGACTCTTCTGAAG 5526
Qy	1596 GCCAACTTAATCTTTAAACGGTAGTTTATCATCAGTCAGCTGCAGACATTTAGAAAAA 1655
Db	
Yq	5527 GCCAACTTAATCTTTAAAGCGTAAGTTTATCACTCACTGTCAGTCAGCATTTAGAAAAA 5586
Qy	1656 TCAAAAGCTGCCGTAGCCCGGAGCAAACCTTAGAAACCTTATGAACTTGGCACTTC 1715
Db	
Yq	5587 TCAAAAGCTGCCGTAGCCCGGAGCAAACCTTAGAAACCTTATGAACTTGGCACTTC 5646
Qy	1716 GGTTTTTTTAATAGAGATCAGGAGGAGCAGGCGGAAACAGACAACCGGATTTAAAAA 1775
Db	
Yq	5647 GGTTTTTTTAATAGAGATCAGGAGGAGCAGGCGGAAACAGACAACCGGATTTAAAAA 5706
Qy	1776 AGGCCACCGCTTTAGTCATCACCTCTAGGCAAGTGGACTTTGGAGGCTCTGGAAAAAGGGA 1835
Db	
Yq	5707 AGGCCACCGCTTTAGTCATCACCTCTAGGCAAGTGGACTTTGGAGGCTCTGGAAAAAGGGA 5766

1836	Qy	AAAGCTGGGCAAAATTGAATGCTTAATAGGGCTTGCCTTCAGTCCGGTCTCAAGGACACT	1895
5767	Db	AAAGCTGGGCAAAATTGAATGCCCTAAATAGGGCTTGCCTTCAGTCCGGTCTCAAGGACACT	5826
1896	Qy	TTAAAAGAAGATTCTCAAGGTAGAAGTAAAGCCGCCCTTCGTCCATGCCCCCTTAATTCAAG	1955
5827	Db	TTAAAAGAAGATTGTCAAGTAGAAGTAGAAGCCGCCCTTCGTCCATGCCCCCTTAATTCAAG	5886
1956	Qy	GGAATCACTGGAAGGCCCACTGCCCCAGGGGCAAAAGAGTCTTTTGGAGTCAGAAGCCACTA	2015
5887	Db	GGAATCACTGGAAGGCCCACTGCCCCAGGGGCAAAAGTCTTTTGGAGTCAGAAGCCACTA	5946
2016	Qy	ACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCCCATCCATGCCATCA	2075
5947	Db	ACCAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGCCCATCCATGCCATCA	6006
2076	Qy	CCCTCACAGAGCCCTGGGTATGCTTGCCANTCAGGGCCAGGAAGTTGTCCTCTCGACA	2135
6007	Db	CCCTCACAGAGCCCTGGGTATGCTTGCCANTCAGGGCCAGGAAGTTGTCCTCTCGACA	6066
2136	Qy	CTGGTGGGTCTTCTTAGTCTTACTCTTCTGTCCCGGCAAACTGTCTCTCAGATCTGTCA	2195
6067	Db	CTGGTGGGTCTTCTTAGTCTTACTCTTCTGTCCCGGCAAACTGTCTCTCAGATCTGTCA	6126
2196	Qy	CTATTCTGAGGGGGTTCNTAAGAAGGGCAGTCACTAGATACCTTTTCCAGCCACTAAAGT	2255
6127	Db	CTATTCTGAGGGGGTCC - TAAGAAGGGCAGTCACTAGATACCTTTTCCAGCCACTAAAGT	6185
2256	Qy	TATGACTGGGGAGCTTTATCTTTTTCACATGCTTTCTTAATATGCTTGAAAGCCCCAC	2315
6186	Db	TATGAACTGGGGAGCTTTATCTTTTTCACATGCTTTCTTAATATGCTTGAAAGCCCCAC	6245
2316	Qy	TACCTTTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACCTGGAACATAGG	2375
6246	Db	TA CCTTGTTAGGGAGAGACATTCTAGCAAAAGCAGGGGCCATTATACCTGGAACATAGG	6305
2376	Qy	AGAAAGAACACCCGTTTGTGTNCCCTGCTGAGGAAGGAATTAATCTCGAAGTCTGGG	2435
6306	Db	AGAAGGAACACCCGTTTGTGTGT - CCCCCTGCTGAGGAAGGAATTAATCTCGAAGTCTGGG	6364
2436	Qy	CAACAGAAGCACAAATATGGAACAGACCAAGAAATGCCGTCTCAAGTTAAACTAAAG	2495
6365	Db	CAACAGAAGCACAAATATGGAACAGACCAAGAAATGCCGTCTCTGTTCAAGTTAAACTAAAG	6424
2496	Qy	GATTCCACTTCTTTTCCCTACCAAGGCGAGTACCCCTCAGACCCCAAGGCCCAACAGGA	2555
6425	Db	GATTCCACTTCTTTTCCCTACCAAGGCGAGTACCCCTCAGACCCCAAGGCCCAACAGGA	6484
2556	Qy	TTCCAAAAGAATTGTTAAGCACTTAAAGCCCAAGGCTTAGTAAAAACATGCATAACTCCC	2615
6485	Db	TTCCAAAAGAATTGTTAAGCACTTAAAGCCCAAGGCTTAGTAAAAACATGCATAACTCCC	6544
2616	Qy	TGCAGTAAATCCGTAGTGGATTGAGGAGGCACAGAACCCAGTGGACAGTGGAGGGTTAG	2675
6545	Db	TGCAGTAAATCCGTAGTGGATTGAGGAGGCACAGAACCCAGTGGACAGTGGAGGGTTAG	6604
2676	Qy	TGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCTCTTTTATACCCAGCTGTACCTAGC	2735
6605	Db	TGCAAGATCTCAGGATTATCAATGGAGGCCGTTGTCTCTTTTATACCCAGCTGTACCTAGC	6664
2736	Qy	CTTTATACTGTG 2747	
6665	Db	CTTTATACTGTG 6676	

RESULT 4  
AA235664  
ID AA235664 standard; cDNA to mRNA; 2372 BP.  
XX  
XX AC AAX25664;  
XX  
XX 21-MAY-1999 (first entry)

XX Human endogenous retrovirus W clone cl.44.4.  
DE  
XX  
XX  
KW Clone; human endogenous retrovirus; genome; autoimmune disease;  
KW multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;  
KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.  
XX  
OS Human endogenous retrovirus.  
XX  
XX W09902696-A1.  
XX  
XX 21-JAN-1999.  
XX  
XX 06-JUL-1998; 98WO-FR001442.  
XX  
XX 07-JUL-1997; 97FR-00008815.  
XX  
XX (INMR ) BIO MERIEUX.  
XX  
XX Beseme F, Blond J, Bouton O, Mandrand B, Mallet F;  
XX WPI; 1999-120897/10.  
XX  
XX New nucleic acid sequences from human endogenous retrovirus-W - expressed  
PT exclusively in placenta and useful in diagnosis and therapy of autoimmune  
PT disease, and abnormal or failed pregnancy.  
XX  
XX Claim 1; Page 68-70; 106pp; French.  
XX  
XX This sequence represents clone cl.44.4 of the human endogenous retrovirus  
CC (HERV) W genome. The nucleic acids, their fragments or peptides encoded  
CC by them are markers of autoimmune disease (e.g. multiple sclerosis,  
CC rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-  
CC dependent diabetes and related pathologies) and of abnormal or  
CC unsuccessful pregnancy and can be used as chromosomal markers for  
CC susceptibility to these conditions, or proximity markers of genes  
CC associated with this susceptibility  
XX  
XX Sequence 2372 BP; 674 A; 593 C; 544 G; 557 T; 0 U; 4 Other;  
SQ

Query Match 11.7%; Score 888; DB 2; Length 2372;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 988; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1507 GGAAGGGAAGAAATCCTACTGCTTCTGGAGAGACTAAGGAGGCAATTGAGGAAGCG 1566  
DB 1383 GGAAGGGAAGAAATCCTACTGCTTCTGGAGAGACTAAGGAGGCAATTGAGGAAGCG 1442  
QY 1567 TGCCTCTCTGACCTGACTCTTCTGAGGCCAACTAATCTTAAAGCGTAAGTTATCAC 1626  
DB 1443 TGCCTCTCTGACCTGACTCTTCTGAGGCCAACTAATCTTAAAGCGTAAGTTATCAC 1502  
QY 1627 TCAGTCAGCTGCAGACATAGAAAACCTTCAAAAGTCTGCCGTAGGCCGGAGCAAAAC 1686  
DB 1503 TCAGTCAGCTGCAGACATAGAAAACCTTCAAAAGTCTGCCGTAGGCCGGAGCAAAAC 1562  
QY 1687 TTAGAAACCTATTTGAACCTTGGCAACVTCGTTTTTTTAAATAGAGATCAGGAGGAGCAG 1746  
DB 1563 TTAGAAACCTATTTGAACCTTGGCAACCTCGTTTTTTTAAATAGAGATCAGGAGGAGCAG 1622  
QY 1747 GCGGAACAGACAAACCGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCTCAGGCA 1806  
DB 1623 GCGGAACAGACAAACCGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCTCAGGCA 1682  
QY 1807 AGTGGACTTTGGAGGCTCTGAAAAGGGAAGCTGGCAAAATTTGAATGCTTAATAGGCC 1866  
DB 1683 AGTGGACTTTGGAGGCTCTGAAAAGGGAAGCTGGCAAAATTTGAATGCTTAATAGGCC 1742  
QY 1867 TTGCTTCCAGTCGGTCTACAAGGACACTTTTAAAAAGATTGTCCAAGTAGAAGTAAGCC 1926  
DB 1743 TTGCTTCCAGTCGGTCTACAAGGACACTTTTAAAAAGATTGTCCAAGTAGAAGTAAGCC 1802  
QY 1927 GCCCCTTCGTCCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCCAGGGG 1986

Db 1803 GCCCCTTCGTCCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGGGG 1852  
QY 1987 ACAAGGTCTTTTGGAGTCAGAAGCCACTAACAGATGATCCAGCAGCAGCACTGAGGGTG 2046  
DB 1863 ACAAGGTCTTTTGGAGTCAGAAGCCACTAACAGATGATCCAGCAGCAGCACTGAGGGTG 1922  
QY 2047 CTTGGGGCAAGCGCCATCCCATGCCATCACCTCACAGAGCCCTGGGTATGCTTACCAT 2106  
DB 1923 CTTGGGGCAAGCGCCATCCCATGCCATCACCTCACAGAGCCCTGGGTATGCTTACCAT 1982  
QY 2107 TGAGGGCCAGGAAGGTCTCTCGGACACTGCTGGGTCTTCTTAGTCTTACTTCTCTG 2166  
DB 1983 TGAGGGCCAGGAAGGTCTCTCGGACACTGCTGGGTCTTCTTAGTCTTACTTCTCTG 2042  
QY 2167 TCCCGGACAACTGCTCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGAGCGGCAGT 2226  
DB 2043 TCCCGGACAACTGCTCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGAGCGGCAGT 2102  
QY 2227 CACTAGATCTTTTCCAGCCACTAAGTTATGAACCTGGGAGCTTTATTTTTCACAT 2286  
DB 2103 CACTAGATCTTTTCCAGCCACTAAGTTATGAACCTGGGAGCTTTATTTTTCACAT 2162  
QY 2287 GCTTTTCTAATATGCTTGAAGCCCACTACCTGCTTAGGGAGACATCTTAGCAAAA 2346  
DB 2163 GCTTTTCTAATATGCTTGAAGCCCACTACCTGCTTAGGGAGACATCTTAGCAAAA 2222  
QY 2347 GCAGGGGCCATTTATACACCTGAACTAGGAGAGGAACACCCCTTTTGTGTCNCCCTGCT 2406  
DB 2223 GCAGGGGCCATTTATACACCTGAACTAGGAGAGGAACACCCCTTTTGTGTCNCCCTGCT 2282  
QY 2407 TGAGGAAGGAATTAATCTTGAAGTCTGGCAACAGAGGACAAATATGGACGACCAAGA 2466  
DB 2283 TGAGGAAGGAATTAATCTTGAAGTCTGGCAACAGAGGACAAATATGGACGACCAAGA 2342  
QY 2467 ATGCCCGCTCTGTTCAAGTTAACTAAAGG 2496  
DB 2343 ATGCCCGCTCTGTTCAAGTTAACTAAAGG 2372

RESULT 5  
AAA59214  
ID AAA59214 standard; DNA; 2372 BP.  
XX  
XX AAA59214;  
XX AC  
XX 07-NOV-2000 (first entry)  
XX DT  
XX  
DE R-U5 region and partial pol gene sequences of HERV-W from human genome.  
XX  
XX Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;  
KW gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.  
XX  
XX Homo sapiens.  
XX OS  
XX W0200043521-A2.  
XX PN  
XX 27-JUL-2000.  
XX PD  
XX  
XX 21-JAN-2000; 2000WO-FR000144.  
XX PF  
XX 21-JAN-1999; 99FR-00000888.  
XX PR  
XX (INMR ) BIO MERIEUX.  
XX PA  
XX Paranhos-Baccala G, Mallet F, Voisset C;  
XX PI  
XX WPI; 2000-499229/44.  
XX DR  
XX New nucleic acid from human endogenous retrovirus, useful e.g. for  
PT diagnosis of autoimmune disease and complications of pregnancy, contains  
PT at least part of the gag gene.  
XX  
XX







151 AGGCTTCCCTGCTAGGACAGAAAGGCCCAAGAGGTAATAAAGGCACTAGTTTCATGAAA 210  
4862 TAATTCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGAGCCCTGCTTTCCAGG 4921  
211 TAATTCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGAGCCCTGCTTTCCAGG 270  
4922 CCACAGTAACCCAGGAGTATCCAGAGCTTTAGGTATACGATATCACTTACACTGCGCCT 4981  
271 CCACAGTAACCCAGGAGTATCCAGAGCTTTAGGTATACGATATCACTTACACTGCGCCT 330  
4982 GAAGGCCACAGTCTTCAGGGAAGGTTCGAGAAATGAATGAAAVCACTCAAAAGGACATCTAA 5041  
331 GAAGGCCACAGTCTTCAGGGAAGGTTCGAGAAATGAATGAAACACTCAAAAGGACATCTAA 390  
5042 AAAAGCAAAACCCAGGAACCCACTCATCGGCTGTCTGTGGCTATAGGCTTTAAAAA 5101  
391 AAAAGCAAAACCCAGGAACCCACTCATCGGCTGTCTGTGGCTATAGGCTTTAAAAA 450  
5102 GAATCTGCAACTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAGGC 5161  
451 GAATCTGCAACTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAGGC 510  
5162 CTTTCATAACCAATGACCTTGTGTGACCCCAAGACAGCCAACTTAGTTGCAGACATCAC 5221  
511 CTTTCATAACCAATGACCTTGTGTGACCCCAAGACAGCCAACTTAGTTGCAGACATCAC 570  
5222 CTCCTTAGCCAAATATCAACAGTTCTTAAAAATTAACAAGAACCTTATCCCTGAGAAGA 5281  
571 CTCCTTAGCCAAATATCAACAGTTCTTAAAAATTAACAAGAACCTTATCCCTGAGAAGA 630  
5282 GGGAAAAGAACTATCCACCCWGTGACATGTTATAGTCAAGTCCCTTCCTTAATTC 5341  
631 GGGAAAAGAACTATCCACCCWGTGACATGTTATAGTCAAGTCCCTTCCTTAATTC 690  
5342 CCCATCCCTAGATACATCTCTGGGAAGGACCTTACCACAGTCAATTTATATACCCCACTGC 5401  
691 CCCATCCCTAGATACATCTCTGGGAAGGACCTTACCACAGTCAATTTATACCCCACTGC 750  
5402 GGTAAAGTGGCTGGAGTGGAGCTTGGATACATCACTTGTAGTCAAAATCTCTGGATCT 5461  
751 GGTAAAGTGGCTGGAGTGGAGCTTGGATACATCACTTGTAGTCAAAATCTCTGGATCT 810  
5462 GCCAAGGAACTGAAATCCAGGAGACACGCTAGCTATTCTCTGTGAACCTCTAGAGGA 5521  
811 GCCAAGGAACTGAAATCCAGGAGACACGCTAGCTATTCTCTGTGAACCTCTAGAGGA 870  
5522 TTTGCGCTGCTCTCTCAAAACAAACACAGGAGAAAGTAACTAAATCAATAATCCCCCA 5581  
871 TTTGCGCTGCTCTCTCAAAACAAACACAGGAGAAAGTAACTAAATCAATAATCCCCCA 929  
5582 TGGSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCCCTSTTTTCACTCTCACTG 5641  
930 TGGSCCTCCCTTATCATATTTTCTCTTTACTGTTCTTTTACCCCTCTTTTCACTCTCACTG 989  
5642 CACCCCTCCATCGCTGTATACAGCATAGTCTCCCTTACCMAGAGTTTCTATGAGAA 5701  
990 CACCCCTCCATCGCTGTATACAGCATAGTCTCCCTTACCMAGAGTTTCTATGAGAA 1049  
5702 TGCAGGCTCCGGAATATTTGATGCCCATCGCTATAGGAGTCTTSTAGGGAACCCCA 5761  
1050 TGCAGGCTCCGGAATATTTGATGCCCATCGCTATAGGAGTCTTCTTAAGGGGAACCCCA 1109  
5762 CTTTCACTGCCACACCCATATGCCCGCACTGCTATCACTCTGCACTCTTTGATGC 5821  
1110 CTTTCACTGCCACACCCATATGCCCGCACTGCTATCACTCTGCACTCTTTGATGC 1169  
5822 ATGCAAAATCTCAATTTGAGACAGGAAATGATTAATCTAGTTGCTCTGGAGCACTTG 5881  
1170 ATGCAAAATCTCAATTTGAGACAGGAAATGATTAATCTAGTTGCTCTGGAGCACTTG 1229  
5882 GAGTCACTGCTCTGAGCACTTACCTACCCAACTGGTATGCTGTATGGGGTGGAGTTC 5941  
1230 GAGTCACTGCTCTGAGCACTTACCTACCCAACTGGTATGCTGTATGGGGTGGAGTTC 1289

QY 5942 AAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCAACTCACCSSGGGTACATG 6001  
Db 1290 AAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCAACTCACCSSGGGTACATG 1349  
QY 6002 GCACCTCTAGCCCTCACAAGGACTAGTCTCTCAAACTACATGAAACCTCGGTACCC 6061  
Db 1350 GCACCTCTAGCCCTCACAAGGACTAGTCTCTCAAACTACATGAAACCTCGGTACCC 1409  
QY 6062 ATACTCGCCTGTPAAGCCTATTTAATACCAACCTCACTGGGCTCCATGAGGTCTCGGCC 6121  
Db 1410 ATACTCGCCTGTPAAGCCTATTTAATACCAACCTCACTGGGCTCCATGAGGTCTCGGCC 1469  
QY 6122 AAAACCTACTAATCTGTGGATATGCTCTCCCTGAACTTTCARGCCATATGTTTCAATTC 6181  
Db 1470 AAAACCTACTAATCTGTGGATATGCTCTCCCTGAACTTTCARGCCATATGTTTCAATTC 1529  
QY 6182 CTGTACTCTGAACAAATGGAACAACTTCAGGACAGAAATAAACACACTTCCTGTTTATAGTAG 6241  
Db 1530 CTGTACTCTGAACAAATGGAACAACTTCAGGACAGAAATAAACACACTTCCTGTTTATAGTAG 1589  
QY 6242 GACCTCTGTTTCCAAATSTGGAATAAACCATACCTCAAACTCCTCACTGTGTAAAAATTTA 6301  
Db 1590 GACCTCTGTTTCCAAATCTGGAATAAACCATACCTCAAACTCCTCACTGTGTAAAAATTTA 1649  
QY 6302 GCAATACATACACAAACAACTCCCAATGCAATCAAGTGGGTAATCTCTCCACACAAA 6361  
Db 1650 GCAATACATACACAAACAACTCCCAATGCAATCAAGTGGGTAATCTCTCCACACAAA 1709  
QY 6362 TAGTCTGCTTACCTCAGGAATATTTTGTCTGTGGTACCTCAGGCTATCGTTGTTTGA 6421  
Db 1710 TAGTCTGCTTACCTCAGGAATATTTTGTCTGTGGTACCTCAGGCTATCGTTGTTTGA 1769  
QY 6422 ATGGCTCTTCAGAACTATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6481  
Db 1770 ATGGCTCTTCAGAACTATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1829  
QY 6482 CTGAACAAGATTTATACAGTTATGTCTATCTAAGCCCGCAACAAAGAGTACCCATTC 6541  
Db 1830 CTGAACAAGATTTATACAGTTATGTCTATCTAAGCCCGCAACAAAGAGTACCCATTC 1889  
QY 6542 TTCTCTTTCTTTATAGAGCAGGAGTCTAGGTGCATCTAGTCTGCGGTATCA 6601  
Db 1890 TTCTCTTTCTTTATAGAGCAGGAGTCTAGGTGCATCTAGTCTGCGGTATCA 1949  
QY 6602 CAACTCTACTCAGTTCTACTCAAACTATCTCAAGAACTAAATGGGACATGGAACGGG 6661  
Db 1950 CAACTCTACTCAGTTCTACTCAAACTATCTCAAGAACTAAATGGGACATGGAACGGG 2009  
QY 6662 TCGCGACTCCCTGGTCACTTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTTC 6721  
Db 2010 TCGCGACTCCCTGGTCACTTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTTC 2069

## RESULT 7

ADS30988

ID ADS30988 standard; DNA; 8523 BP.

XX ADS30988;

AC ADS30988;

XX 18-NOV-2004 (first entry)

XX Human genome high complexity repeat found in the HIRA gene #21.

XX Human; ds;

XX histone cell cycle regulation defective, *S. cerevisiae* homologue A; HIRA;

XX high complexity repeat; in situ hybridisation; Southern blot;

XX chromosome breakpoint; inherited genetic disease; neoplastic disorder;

XX chromosome 22; Digorge syndrome; Velo-Cardio-facial syndrome.

XX Homo sapiens.

XX US2003224356-A1.

PN

XX PD 04-DEC-2003.

XX PF 14-MAY-2001; 2001US-00854867.

XX PR 16-MAY-2000; 2000US-00573080.

XX PA (KNOLL/) KNOLL J H M.

XX PA (ROGA/) ROGAN P K.

XX PI Knoll JHM, Rogan PK;

XX DR WPI; 2002-062378/08.

XX PT Single copy genomic hybridization probes for detecting specific nucleic acid sequences in sample by in situ hybridization useful for detection of acquired or inherited genetic diseases.

XX PS Example 1; SEQ ID NO 21; 30pp; English.

XX CC The invention relates to a nucleic acid hybridisation probe comprising a labelled, single copy nucleic acid of at least 50 nucleotides, which will hybridise to a deduced single copy sequence interval in target nucleic acid (TNA) of known sequence. The single copy sequence is deduced by comparing the target nucleic acid (e.g. a disease causing gene) with a collection of high and low complexity repeat sequences as found in the genome of the organism from containing the target nucleic acid. The probe is generated by PCR on the target sequence. The probe is essentially free of blocking nucleic acid sequences which will hybridise to repeat sequences within the genome of which the TNA is a part, and is labelled with a label selected from fluorochrome-responsive labels, fluorochromes, calorimetric chemical, conjugated proteins, antibodies, antigens and their mixtures. The probe is useful in a hybridisation method, where the hybridisation method is from in situ hybridisation, Southern blot, and other methods in which nucleic acid is immobilised, where the method further comprises selecting a single copy nucleic acid which will hybridise to a duplication or triplicon sequence domain. The probe is useful for determining the existence of previously unknown repeat sequence families in a genome. The method comprises reacting a labelled probe with the genome, causing the probe to hybridise and ascertaining if the probe hybridises to the genome at more than three preferably ten different locations as a determination of new repeat sequence family, where the determining step comprises selecting the single copy sequence from a duplication or triplicon sequence domain. The probe is useful for determining a chromosome breakpoint and is useful in the fields for cytogenetics and molecular genetics for determining the presence of specific nucleic acid sequences in a sample of eukaryotic origin, e.g. the probes may be used to analyse specific chromosomal locations by in situ hybridisation as a detection of acquired or inherited genetic diseases especially for detection of genetic or neoplastic disorders. Unlike prior art techniques, the probe permits more precise chromosomal breakpoint determinations by in situ hybridisation. The genomic sequence comprising the human HIRA gene (histone cell cycle regulation defective, S. cerevisiae, homologue A) was analysed for single copy sequence intervals for use as probes of the invention. HIRA is located on chromosome 22 as a duplicate, deletions of 1 copy lead to DiGeorge and Velo-Cardio-facial syndromes. The present sequence is a high complexity repeat found within the human genome used to analyse the HIRA gene for repeat regions. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at [seqdata.uspto.gov/sequence.html?DocID=20030224356](http://seqdata.uspto.gov/sequence.html?DocID=20030224356).

XX SQ Sequence 8523 BP; 2560 A; 2118 C; 1814 G; 2031 T; 0 U; 0 Other;

Query Match 10.3%; Score 779; DB 7; Length 8523;

Best Local Similarity 98.9%; Pred. No. 0;

Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 4802 AGGCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCATGAAA 4861

DB 6071 AGGCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCATGAAA 6130

QY 4862 TAAATCCAGATTCCGACTTCCCGAGGGCTTACAGAGTGCACATAGCCCTCTTCCAGG 4921

Db 6131 TAATTCAGATTCCGACTTCCCGAGGGCTTACAGAGTGCACATAGCCCTCTTCCAGG 6190

QY 4922 CCACAGTAACCCAGGAGTATCCAGGCTTAGGTATAGGTATACATTACACTCGGCT 4981

Db 6191 CCACAGTAACCCAGGAGTATCCAGGCTTAGGTATAGGTATACATTACACTCGGCT 6250

QY 4982 GAAGGCCACACTCTCCAGGAGGTTCGAGAAAATGAATGAAYACTCAAGAGACATCTAA 5041

Db 6251 GAAGGCCACACTCTCCAGGAGGTTCGAGAAAATGAATGAAYACTCAAGAGACATCTAA 6310

QY 5042 AAAAGCAAAACCCAGGAAACCCACCTCAATGGCCCTGTTCTGTCTATAGCTTTAAAAA 5101

Db 6311 AAAAGCAAAACCCAGGAAACCCACCTCAATGGCCCTGTTCTGTCTATAGCTTTAAAAA 6370

QY 5102 GAATCTGCACTTTCCCAAAAAGAGACTTAGCCCATAGCAATGCTGTATGAAGGC 5161

Db 6371 GAATCTGCACTTTCCCAAAAAGAGACTTAGCCCATAGCAATGCTGTATGAAGGC 6430

QY 5162 CCTTCATAACCAATGACTTGTCTGTACCCCAAGACAGCAACTTAGTTGACAGACATCAC 5221

Db 6431 CCTTCATAACCAATGACTTGTCTGTACCCCAAGACAGCAACTTAGTTGACAGACATCAC 6490

QY 5222 CTCTCTAGCCAAATATCAACAAGTTCTTTAAAAACATTACAAGGAACCTATCCCTGAGAAGA 5281

Db 6491 CTCTCTAGCCAAATATCAACAAGTTCTTTAAAAACATTACAAGGAACCTATCCCTGAGAAGA 6550

QY 5282 GGGAAAAGAACTATTCACACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTTAATTC 5341

Db 6551 GGGAAAAGAACTATTCACACCCCTTGTGACATGGTATTAGTCAAGTCCCTTCYCTTAATTC 6610

QY 5342 CCCATCCCTAGATACATCCTCGGAGGACCTACCCAGTCAATTTATATACCCCAACTGC 5401

Db 6611 CCCATCCCTAGATACATCCTCGGAGGACCTACCCAGTCAATTTATATACCCCAACTGC 6670

QY 5402 GGTAAAGTGGCTGGAGTCTTGATATACATCACACTTGAGTCAAAATCCTGGATACT 5461

Db 6671 GGTAAAGTGGCTGGAGTCTTGATATACATCACACTTGAGTCAAAATCCTGGATACT 6730

QY 5462 GCCAAGGAACCTGAAAATCCAGGAGACAAAGCTAGTATTCTGTGAACCTCTAGAGA 5521

Db 6731 GCCAAGGAACCTGAAAATCCAGGAGACAAAGCTAGTATTCTGTGAACCTCTAGAGA 6790

QY 5522 TTTGGCCCTGCTCTTCAACACACACAGGAGGAAGTAACTAAATCATAAATCCCCA 5581

Db 6791 TTTGGCCCTGCTCTTCAACACACACAGGAGGAAGTAACTAAATCATAAATCCCCA 6849

QY 5582 TGGSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTACCTSTTTTCACTCTCCTG 5641

Db 6850 TGGSCCTCCCTTATCATATTTTCTCTTACTGTTCTTTTACCTCTTCTCACTCTCCTG 6909

QY 5642 CACCCCTCCATGCGCTGTATGACAGTAGTCCCTTACCMAGAGTTTCTATGAGAAA 5701

Db 6910 CACCCCTCCATGCGCTGTATGACAGTAGTCCCTTACCMAGAGTTTCTATGAGAAA 6969

QY 5702 TGCAGCTCCCGAAATTTGATGCCCATCGTATAGAGTCTTTSTAAGGGAACCCCA 5761

Db 6970 TGCAGCTCCCGAAATTTGATGCCCATCGTATAGAGTCTTTSTAAGGGAACCCCA 7029

QY 5762 CTTTCACTGCCACACACCATATGCCCCGCACTGCTATCCTGCTCTGTCCTTTTGATGC 5821

Db 7030 CTTTCACTGCCACACACCATATGCCCCGCACTGCTATCCTGCTCTGTCCTTTTGATGC 7089

QY 5822 ATGCAAAATATCATTTATTGGACAGAAAATGATTAATCCTAGTTGTCTCGAGGACTTGG 5881

Db 7090 ATGCAAAATATCATTTATTGGACAGAAAATGATTAATCCTAGTTGTCTCGAGGACTTGG 7149

QY 5882 GAGTCACTGCTGTTGGACTTACTTCAACCAACTGGTATGCTCATGTTGGGGTGGAGTTTC 5941

Db 7150 GAGTCACTGCTGTTGGACTTACTTCAACCAACTGGTATGCTCATGTTGGGGTGGAGTTTC 7209

QY 5942 AAGATCAGGCAAGAGAAAACATGTAAAGAAAGTAAATCTCCCAACTCACCSGGGTACATG 6001

Db 7210 AAGATCAGGCAAGAGAAAAACATCTAAGAGAGTAATCTCCCACTCACCGGGTACATG 7269  
QY 6002 GCACCTTACGCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCTCGTACCC 6061  
Db 7270 GCACCTTACGCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCTCGTACCC 7329  
QY 6062 ATACTCGCCTGGTAAGCCTATTTAATACACCCCTCAGCTGGGCTCCATGAGGTCTCGGCC 6121  
Db 7330 ATACTCGCCTGGTAAGCCTATTTAATACACCCCTCAGCTGGGCTCCATGAGGTCTCGGCC 7389  
QY 6122 AAAACCTACTAATGTTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATCC 6181  
Db 7390 AAAACCTACTAATGTTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATCC 7449  
QY 6182 CTGTACTGACCAATGGAACAACTTCAGACAGAAATAAACAACATTCCTCGTTTGTAGTAG 6241  
Db 7450 CTGTACTGACCAATGGAACAACTTCAGACAGAAATAAACAACATTCCTCGTTTGTAGTAG 7509  
QY 6242 GACCTCTGTTTCAATSTGGAATAACCCATACCTCAAACTCAGCTGGTGAATAATTTA 6301  
Db 7510 GACCTCTGTTTCAATSTGGAATAACCCATACCTCAAACTCAGCTGGTGAATAATTTA 7569  
QY 6302 GCAATACTACATACACAACTCCCAATGTCATCAGTGGGTAACTCTCTCCACACAAA 6361  
Db 7570 GCAATACTACATACACAACTCCCAATGTCATCAGTGGGTAACTCTCTCCACACAAA 7629  
QY 6362 TAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTGACTCAGCTATCGTGTGTGA 6421  
Db 7630 TAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTGACTCAGCTATCGTGTGTGA 7689  
QY 6422 ATGCTCTTCAGAACTATGCTCTCTCTCATCTTAGTGCCTCCATGRCATCTACA 6481  
Db 7690 ATGCTCTTCAGAACTATGCTCTCTCTCATCTTAGTGCCTCCATGRCATCTACA 7749  
QY 6482 CTGAACAAGATTATACAGTTATGTCTATATCTAAGCCCGCAACAAAGAGTACCCATTC 6541  
Db 7750 CTGAACAAGATTATACAGTTATGTCTATATCTAAGCCCGCAACAAAGAGTACCCATTC 7809  
QY 6542 TTCCTTTTGTATAGGACGAGGAGTGTAGGTGACCTAGGTACTGGCATTGGCGGTATCA 6601  
Db 7810 TTCCTTTTGTATAGGACGAGGAGTGTAGGTGACCTAGGTACTGGCATTGGCGGTATCA 7869  
QY 6602 CACCTCTACTAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGG 6661  
Db 7870 CACCTCTACTAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGG 7929  
QY 6662 TCGCCGACTCCCTGGTCACTTCGAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTC 6721  
Db 7930 TCGCCGACTCCCTGGTCACTTCGAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTC 7989

## RESULT 8

ADV36376

ID ADV36376 standard; DNA; 8523 BP.

AC ADV36376;

XX

05-MAY-2005 (first entry)

HIRA genomic fragment SEQ ID NO 21.

XX

XX hybridization; DNA detection; neoplasm; genetic disorder; cytogenetics;  
KW HIRA; ds.

XX Homo sapiens.

XX

XX W0200188089-A2.

XX

XX 22-NOV-2001.

XX

XX 15-MAY-2001; 2001WO-US015674.

XX

XX 16-MAY-2000; 2000US-00573080.

XX

XX

XX

XX

XX

XX

XX

PR 14-MAY-2001; 2001US-00854867.

XX (CHIL-) CHILDREN'S MERCY HOSPITAL.

XX

XX Knoll JHM, Rogan PK, Cazarro PM;

XX

XX WPI; 2002-062378/08.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

SQ Sequence 8523 BP; 2560 A; 2118 C; 1814 G; 2031 T; 0 U; 0 Other;

Query Match 10.3%; Score 779; DB 7; Length 8523;

Best Local Similarity 98.9%; Pred. No. 0;

Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 4802 AGGCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTATAAAGGCATAGTTTCATGAAA 4861

Db 6071 AGGCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTATAAAGGCATAGTTTCATGAAA 6130

QY 4862 TAATTCCTCAGATTCCGACTTCCCGAGGCTTACAGAGTACATAGCTTCTTCAGG 4921

Db 6131 TAATTCCTCAGATTCCGACTTCCCGAGGCTTACAGAGTACATAGCTTCTTCAGG 6190

QY 4922 CCACAGTAACCCAGGAGGTATCCAGGCGTTAGGTATACGATATCACTTACATCTGGCGCT 4981

Db 6191 CCACAGTAACCCAGGAGGTATCCAGGCGTTAGGTATACGATATCACTTACATCTGGCGCT 6250

QY 4982 GAAGGCCACAGTCTTCAGGGAAGGTCGAGAAATGAATGAAAYACTCAAGGACATCTAA 5041

Db 6251 GAAGGCCACAGTCTTCAGGGAAGGTCGAGAAATGAATGAAAYACTCAAGGACATCTAA 6310

QY 5042 AAAAGCAAAACCAGGAAACCCACCTCACATGCGCTGTCTGTTGCTTATAGCTTAAAAA 5101

Db 6311 AAAAGCAAAACCAGGAAACCCACCTCACATGCGCTGTCTGTTGCTTATAGCTTAAAAA 6370

```
QY 5102 GAATCTGCAACTTTCCCAAAAGCAGGACTTAGCCCATAGAAATGCTGTATGAAGC 5161
DB 6371 GAATCTGCAACTTTCCCAAAAGCAGGACTTAGCCCATAGAAATGCTGTATGAAGC 6430
QY 5162 CTTTCATAACCAATGACCTTGTGCTTGAGCCCAAGCAGCACAATCTAGTTGTCAGACATCAC 5221
DB 6431 CCTTCATAACCAATGACCTTGTGCTTGAGCCCAAGCAGCACAATCTAGTTGTCAGACATCAC 6490
QY 5222 CTCCTTAGCCAAATATCAACAAGTTCTTAAACATTAACAGAACTATCCCTGAGAAGA 5281
DB 6491 CTCCTTAGCCAAATATCAACAAGTTCTTAAACATTAACAGAACTATCCCTGAGAAGA 6550
QY 5282 GGGAAAAGAACTATTCCACCCWGGGATGATGTTAGTCAAGTCCCTTCYCTAAATTC 5341
DB 6551 GGGAAAAGAACTATTCCACCCCTTGTGATAGTGTATAGTCAAGTCCCTTCCTCTAAATTC 6610
QY 5342 CCCATCCCTTAGATACATCTCGGGAAGGACCCCTACCCAGTCAATTTATATACCCCAACTGC 5401
DB 6611 CCCATCCCTTAGATACATCTCGGGAAGGACCCCTACCCAGTCAATTTATCTACCCCAACTGC 6670
QY 5402 GGTAAAAGTGGCTGAGTGGAGTCTTGGATACATCACACTTGGAGTCAAAATCTCGATACT 5461
DB 6671 GGTAAAAGTGGCTGAGTGGAGTCTTGGATACATCACACTTGGAGTCAAAATCTCGATACT 6730
QY 5462 GCCAAAGAACCTGAAATCCAGGAGACAAAGCTAGCTATTCTGTGAACCTCTAGAGA 5521
DB 6731 GCCAAAGAACCTGAAATCCAGGAGACAAAGCTAGCTATTCTGTGAACCTCTAGAGA 6790
QY 5522 TTGTGGCCCTGCTCTTCAAAACAACACAGGAGGAAAGTAACTAAATCATAAATCCCCCA 5581
DB 6791 TTGTGGCCCTGCTCTTCAAAACAACACAGGAGGAAAGTAACTAAATCATAAATCCCCCA 6849
QY 5582 TGGSCCTCCCTTATCATATTTTCTCTKPASTGTTSTTTACCTSTTTCACTCTCACTG 5641
DB 6850 TGGSCCTCCCTTATCATATTTTCTCTTACTGTTCTTTACCTCTTCACTCTCACTG 6909
QY 5642 GACCCCTCCATGCGCTGTATGACCACTAGCTAGCTTCCCTTACCMAGAGTTTCTATGGANA 5701
DB 6910 GACCCCTCCATGCGCTGTATGACCACTAGCTAGCTTCCCTTACCMAGAGTTTCTATGGANA 6969
QY 5702 TGCAGCGTCCCGGAATATTGATGCCCCCATCTATAGGAGTCTTTSTAAGGGAACCCCCA 5761
DB 6970 TGCAGCGTCCCGGAATATTGATGCCCCCATCTATAGGAGTCTTTSTAAGGGAACCCCCA 7029
QY 5762 CTTTCACTGCCCCACACCCCATATGCCCCGCAACTGTCTATCACTCTGCACTCTTTGGATGC 5821
DB 7030 CTTTCACTGCCCCACACCCCATATGCCCCGCAACTGTCTATCACTCTGCACTCTTTGGATGC 7089
QY 5822 ATGCAATACTCATTTATTGGACAGGAAATGATTAATCCCTAGTTGTCCTGGAGGACTTG 5881
DB 7090 ATGCAATACTCATTTATTGGACAGGAAATGATTAATCCCTAGTTGTCCTGGAGGACTTG 7149
QY 5882 GAGTCACTCTCTGTTGGACTTACTTCAACCAACTGGTATGCTCATGCGGGGTGGAGTTC 5941
DB 7150 GAGTCACTCTCTGTTGGACTTACTTCAACCAACTGGTATGCTCATGCGGGGTGGAGTTC 7209
QY 5942 AAGATCAGGCAAGAGAAAAACATGTAAAGAGTAACTCTCCCAACTCACCGGGTACATG 6001
DB 7210 AAGATCAGGCAAGAGAAAAACATGTAAAGAGTAACTCTCCCAACTCACCGGGTACATG 7269
QY 6002 GCACCTCTAGCCCTTACAAGAGCTAGATCTCTCAAAACTACATGAACCCCTCCGTACCC 6061
DB 7270 GCACCTCTAGCCCTTACAAGAGCTAGATCTCTCAAAACTACATGAACCCCTCCGTACCC 7329
QY 6062 ATACTCGCCTGTAGCCTATTATTAACCAACCTCACTGGGCTCCATGAGGTCTCGGCC 6121
DB 7330 ATACTCGCCTGTAGCCTATTATTAACCAACCTCACTGGGCTCCATGAGGTCTCGGCC 7389
QY 6122 AAAACCCCTACTACTGTTGGATATGCTCCCTGAACTTTCARGGCATATGTTCAATCC 6181
DB 7390 AAAACCCCTACTACTGTTGGATATGCTCCCTGAACTTTCARGGCATATGTTCAATCC 7449
QY 6182 CTGTACTGAAACAATGGAACAACCTTCAGCACAGAAATTAACACCACCTTCGGTTTTAGTAG 6241
```

```
DB 7450 CTGTACTTGACATGGAACAACCTTCAGCACAGAAATATAACCACTTCGGTTTTAGTAG 7509
QY 6242 GACCTCTTGTGTTTCAATSTGGAATAAACCCATACTCAAACTCACTGTTGTAATAATTTA 6301
DB 7510 GACCTCTTGTGTTTCAATSTGGAATAAACCCATACTCAAACTCACTGTTGTAATAATTTA 7569
QY 6302 GCAATACTACATACACAAACCACTCCCAATGCATCAGGTGGGTAACTCCTCCACACAAA 6361
DB 7570 GCAATACTACATACACAAACCACTCCCAATGCATCAGGTGGGTAACTCCTCCACACAAA 7629
QY 6362 TAGTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCCTATCGTTGTTTGA 6421
DB 7630 TAGTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCCTATCGTTGTTTGA 7689
QY 6422 ATGGCTCTTCAGAAATCTATAGTGTCTCTCATTTAGTGTCCCTATGRCCTATCA 6481
DB 7690 ATGGCTCTTCAGAAATCTATAGTGTCTCTCATTTAGTGTCCCTATGRCCTATCA 7749
QY 6482 CTGAACAAGATTTATACAGTTATGTATATCTAAGCCCCGCAACAAAGAGTACCCATTC 6541
DB 7750 CTGAACAAGATTTATACAGTTATGTATATCTAAGCCCCGCAACAAAGAGTACCCATTC 7809
QY 6542 TTCTCTTTGTTTATAGGAGCAGAGTGTAGGTGCACTAGGTACTGGCATTTGGCGGTATCA 6601
DB 7810 TTCTCTTTGTTTATAGGAGCAGAGTGTAGGTGCACTAGGTACTGGCATTTGGCGGTATCA 7869
QY 6602 CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAAACGG 6661
DB 7870 CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAAACGG 7929
QY 6662 TCGCGGACTCCCTGGTCACTTGAAGATCACTTAACTCCCTAGCAGCAGTAGTCCCTTC 6721
DB 7930 TCGCGGACTCCCTGGTCACTTGAAGATCACTTAACTCCCTAGCAGCAGTAGTCCCTTC 7989

RESULT 9
ABN97929
ID ABN97929 standard; DNA; 10499 BP.
XX AC ABN97929;
XX DT 01-AUG-2002 (first entry)
XX DE Human retroviral sequence HERV-7q.
XX KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
XX KW multiple sclerosis; ds.
XX OS Human endogenous retrovirus.
XX PN WO9967395-A1.
XX PD 29-DEC-1999.
XX PF 23-JUN-1999; 99WO-FR001513.
XX PR 23-JUN-1998; 98FR-00007920.
XX PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX PI Alliel PM, Perin J, Rieger F;
XX DR WPI; 2000-160587/14.
XX PT New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
XX PT for diagnosis, treatment and prevention of autoimmune and neurological
XX PT diseases.
XX PS Claim 3; Fig 1; 225pp; French.
XX CC The present invention relates to new nucleic acid sequences of human
endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
```





RESULT 10  
 ABL61744  
 ID ABL61744 standard; DNA; 56093 BP.  
 XX AC ABL61744;  
 XX DT 15-MAY-2002 (first entry)  
 XX DE Colon adenocarcinoma related gene sequence SEQ ID NO: 81.  
 XX KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
 KW gene; ds.  
 XX OS Homo sapiens.  
 XX PN WO200194629-A2.  
 XX PD 13-DEC-2001.  
 XX PF 30-MAY-2001; 2001WO-US010838.  
 XX PR 05-JUN-2000; 2000US-0209473P.  
 PR 05-JUN-2000; 2000US-0209531P.  
 PR 18-SEP-2000; 2000US-0233133P.  
 PR 20-SEP-2000; 2000US-0234009P.  
 PR 20-SEP-2000; 2000US-0234034P.  
 PR 20-SEP-2000; 2000US-0234052P.  
 PR 22-SEP-2000; 2000US-0234509P.  
 PR 22-SEP-2000; 2000US-0234567P.  
 PR 25-SEP-2000; 2000US-0234923P.  
 PR 25-SEP-2000; 2000US-0234924P.  
 PR 25-SEP-2000; 2000US-0235077P.  
 PR 25-SEP-2000; 2000US-0235082P.  
 PR 25-SEP-2000; 2000US-0235134P.  
 PR 25-SEP-2000; 2000US-0235280P.  
 PR 26-SEP-2000; 2000US-0235637P.  
 PR 26-SEP-2000; 2000US-0235638P.  
 PR 27-SEP-2000; 2000US-0235711P.  
 PR 27-SEP-2000; 2000US-0235720P.  
 PR 27-SEP-2000; 2000US-0235840P.  
 PR 27-SEP-2000; 2000US-0235863P.  
 PR 28-SEP-2000; 2000US-0236028P.  
 PR 28-SEP-2000; 2000US-0236032P.  
 PR 28-SEP-2000; 2000US-0236033P.  
 PR 28-SEP-2000; 2000US-0236034P.  
 PR 28-SEP-2000; 2000US-0236109P.  
 PR 28-SEP-2000; 2000US-0236111P.  
 PR 29-SEP-2000; 2000US-0236842P.  
 PR 29-SEP-2000; 2000US-0236891P.  
 PR 02-OCT-2000; 2000US-0237172P.  
 PR 02-OCT-2000; 2000US-0237173P.  
 PR 02-OCT-2000; 2000US-0237278P.  
 PR 02-OCT-2000; 2000US-0237294P.  
 PR 02-OCT-2000; 2000US-0237295P.  
 PR 02-OCT-2000; 2000US-0237316P.  
 PR 03-OCT-2000; 2000US-0237425P.  
 PR 03-OCT-2000; 2000US-0237598P.  
 PR 03-OCT-2000; 2000US-0237604P.  
 PR 03-OCT-2000; 2000US-0237606P.  
 PR 03-OCT-2000; 2000US-0237608P.  
 PR 01-NOV-2000; 2000US-0244867P.  
 PR 01-NOV-2000; 2000US-0245084P.  
 XX (AVAL-) AVALON PHARM.  
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
 PI Soppet DR, Weaver Z;  
 XX WPI; 2002-188264/24.

XX Screening for anti-neoplastic agent involves exposing cells to a chemical  
 PT agent to be tested for anti-neoplastic activity, and determining a change  
 PT in expression of a gene of a signature gene set.  
 XX Claim 1; SEQ ID NO 81; 44pp; English.  
 XX The present invention describes a method (M1) for screening for an anti-  
 CC neoplastic agent. The method involves exposing cells to a chemical agent  
 CC to be tested for anti-neoplastic activity, determining a change in  
 CC expression of at least one gene (I) of a signature gene set, where (I)  
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
 CC to ABL70110), or is at least 95% identical to (S), where a change in  
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
 CC activity and can be used in gene therapy. M1 can be used for screening an  
 CC anti-neoplastic agent, and can be used for producing a product which is  
 CC the data collected with respect to the anti-neoplastic agent as a result  
 CC of M1, and the data is sufficient to convey the chemical structure and/or  
 CC properties of the agent. M1 can be used in the treatment of cancer such  
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,  
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell  
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous  
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's  
 CC tumour  
 XX Sequence 56093 BP; 16164 A; 12346 C; 10702 G; 16881 T; 0 U; 0 Other;  
 SQ Query Match 10.3%; Score 779; DB 6; Length 56093;  
 Best Local Similarity 98.9%; Pred. No. 0;  
 Matches 1899; Conservative 0; Mismatches 20; Indels 1; Gaps 1;  
 QY 4802 AGGCTTCCCTGTAGGACAGAAAGGCCGAGAGGTAAAGGCATAGTTCATGAAA 4861  
 DB 35101 AGGCTTCCCTGTAGGACAGAAAGGCCGAGAGGTAAAGGCATAGTTCATGAAA 35160  
 QY 4862 TAAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTCCAGG 4921  
 DB 35161 TAAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTCCAGG 35220  
 QY 4922 CCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACATGCGCCT 4981  
 DB 35221 CCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACATGCGCCT 35280  
 QY 4982 GAAGGCCACAGTCTCAGGGAAGTTCGAGAAATGAATGAAYACTCAAGACATCTAA 5041  
 DB 35281 GAAGGCCACAGTCTCAGGGAAGTTCGAGAAATGAATGAAYACTCAAGACATCTAA 35340  
 QY 5042 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCCTGCTGTGCTATAGCCTTAAAAA 5101  
 DB 35341 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCCTGCTGTGCTATAGCCTTAAAAA 35400  
 QY 5102 GAATCTGCAACTTTCCCAAAAGCAGGACTTAGCCCATACCAAAATGCTGTATGGAAGGC 5161  
 DB 35401 GAATCTGCAACTTTCCCAAAAGCAGGACTTAGCCCATACCAAAATGCTGTATGGAAGGC 35460  
 QY 5162 CCTTCATTAACCAATGACCTTGCTTGACCCAGCAGCACTTAGTGTGACAGACATCAC 5221  
 DB 35461 CCTTCATTAACCAATGACCTTGCTTGACCCAGCAGCACTTAGTGTGACAGACATCAC 35520  
 QY 5222 CTCCTTAGCCAAATATCAACAAGTTCTTAAACATTTACAAAGCACTTATCCCTGAGAAGA 5281  
 DB 35521 CTCCTTAGCCAAATATCAACAAGTTCTTAAACATTTACAAAGCACTTATCCCTGAGAAGA 35580  
 QY 5282 GGGAAAAGAACTATTTCACCCWGTGACATGTATAGTCAAGTCCCTTCCTCTAATTC 5341  
 DB 35581 GGGAAAAGAACTATTTCACCCWGTGACATGTATAGTCAAGTCCCTTCCTCTAATTC 35640  
 QY 5342 CCCATCCCTAGATACATCTGGGAGGACCTTACCAGTCATTTATYACCCCACTGC 5401  
 DB 35641 CCCATCCCTAGATACATCTGGGAGGACCTTACCAGTCATTTATYACCCCACTGC 35700  
 QY 5402 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACATCTTGAGTCAAACTCTCGGATACT 5461



```
Db 35701 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAACTCCTGGATCT 35760
Qy 5462 GCCAAGGAACCTGAAATCGAGAGCAACGGTAGCTATTCTGTGAACTCTAGAGGA 5521
Db 35761 GCCAAGGAACCTGAAATCGAGAGCAACGGTAGCTATTCTGTGAACTCTAGAGGA 35820
Qy 5522 TTTCGGCTCTCTTTCAACAACAACACGAGGAGAAAGTAATAAATCAATAATCCCCCA 5581
Db 35821 TTTCGGCTCTCTTTCAACAACAACACGAGGAGAAAGTAATAAATCAATAAT-CCCCA 35879
Qy 5582 TGSCCTCCCTTATCATATTTTCTCTKTAATSTTSTTTTACCCTCTTTTCACTCTCACTG 5641
Db 35880 TGSCCTCCCTTATCATATTTTCTCTTACTTGTCTTTTACCCTCTTTTCACTCTCACTG 35939
Qy 5642 CACCCCTCCATGCGCTGTATGACAGTAGCTCCCTCTACCMAGAGTTTCTATGAGAA 5701
Db 35940 CACCCCTCCATGCGCTGTATGACAGTAGCTCCCTCTACCAAGAGTTTCTATGAGAA 35999
Qy 5702 TGCAGCTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGAACCCCA 5761
Db 36000 TGCAGCTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGAACCCCA 36059
Qy 5762 CTTCACTGCCACACCCATATGATGCCCGCAACTGCTATCACTCTGCCACTCTTTTGCATGC 5821
Db 36060 CTTCACTGCCACACCCATATGATGCCCGCAACTGCTATCACTCTGCCACTCTTTTGCATGC 36119
Qy 5822 ATGCAATACTCATTTATGACAGGAAATGATTAATCTAGTTGCTCGGAGGACTTG 5881
Db 36120 ATGCAATACTCATTTATGACAGGAAATGATTAATCTAGTTGCTCGGAGGACTTG 36179
Qy 5882 GAGTCACTGCTCTGTGACTTATCTTACCCAACTGGTATGCTGATGGGGTGGAGTTC 5941
Db 36180 GAGTCACTGCTCTGTGACTTATCTTACCCAACTGGTATGCTGATGGGGTGGAGTTC 36239
Qy 5942 AAGATCAGGCAAGAGAAATCATGTAAGAAAGTAATCTCCCAACTCACCGGGGTACATG 6001
Db 36240 AAGATCAGGCAAGAGAAATCATGTAAGAAAGTAATCTCCCAACTCACCGGGGTACATG 36299
Qy 6002 GCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCC 6061
Db 36300 GCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCC 36359
Qy 6062 ATACTCGCTGGTAAGCCTATTTAATACACCTCTACTGGGCTCCATGAGGTCTCGGCC 6121
Db 36360 ATACTCGCTGGTAAGCCTATTTAATACACCTCTACTGGGCTCCATGAGGTCTCGGCC 36419
Qy 6122 AAAACCTCTACTACTGTGTGATATGCTCCCTGAACTTCARGCCATATGTTTCAATCC 6181
Db 36420 AAAACCTCTACTACTGTGTGATATGCTCCCTGAACTTCARGCCATATGTTTCAATCC 36479
Qy 6182 CTGTACTGAACTGAAACAACTTCAGCACAGAAATAAACACCACTTCGGTTTTAGTAG 6241
Db 36480 CTGTACTGAACTGAAACAACTTCAGCACAGAAATAAACACCACTTCGGTTTTAGTAG 36539
Qy 6242 GACCTCTGTTTCCATSTGAAATTAACCATACCTCAAACTCACTGCTGTGTAATAATTA 6301
Db 36540 GACCTCTGTTTCCATSTGAAATTAACCATACCTCAAACTCACTGCTGTGTAATAATTA 36599
Qy 6302 GCAATACTACATACAAACCACTCCCAATGCATCAGTGGTAACTCTCCACACAAA 6361
Db 36600 GCAATACTACATACAAACCACTCCCAATGCATCAGTGGTAACTCTCCACACAAA 36659
Qy 6362 TAGTCTGCCTTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCGTTGTTTGA 6421
Db 36660 TAGTCTGCCTTACCTCAGGAATATTTTGTCTGTGTGTACCTCAGCCTATCGTTGTTTGA 36719
Qy 6422 ATGGCTCTCAGAACTATGTGCTCTCTCATTTAGTGGCCCTCCATGRCATCTACA 6481
Db 36720 ATGGCTCTCAGAACTATGTGCTCTCTCTCATTTAGTGGCCCTCTGACCATCTACA 36779
Qy 6482 CTGAACAAGATTTATACAGTTATGTATATCTAAGCCCGCAACAAAGAGTACCATTC 6541
Db 36780 CTGAACAAGATTTATACAGTTATGTATATCTAAGCCCGCAACAAAGAGTACCATTC 36839
```

---

```
Qy 6542 TTCTTTTGTATAGGAGCAGGAGTCTAGTGCACCTAGTACTGGCATTTGGCGGTATCA 6601
Db 36840 TTCTTTTGTATAGGAGCAGGAGTCTAGTGCACCTAGTACTGGCATTTGGCGGTATCA 36899
Qy 6602 CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG 6661
Db 36900 CAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG 36959
Qy 6662 TCGCCGACTCCCTGGTCACCTTCGAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTC 6721
Db 36960 TCGCCGACTCCCTGGTCACCTTCGAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTC 37019

RESULT 11
AAS84210
ID AAS84210 standard; cDNA; 6394 BP.
XX
AC AAS84210;
XX
DT 13-FEB-2002 (first entry)
XX
DNA encoding novel human diagnostic protein #20014.
XX
Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN W0200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
(HYSE-) HYSEQ INC.
XX
Drmanac RT, Liu C, Tang YT;
XX
WPI; 2001-639362/73.
DR P-PSDB; ABG20023.
XX
New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity.
XX
Claim 1; SEQ ID NO 20014; 103pp; English.
XX
The invention relates to isolated polynucleotide (I) and polypeptide (II)
sequences. (I) is useful as hybridisation probes, polymerase chain
reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
and in recombinant production of (II). The polynucleotides are also used
in diagnostics as expressed sequence tags for identifying expressed
genes. (II) is useful in gene therapy techniques to restore normal
activity of (II) or to treat disease states involving (II). (II) is
useful for generating antibodies against it, detecting or quantitating a
polypeptide in tissue, as molecular weight markers and as a food
supplement. (II) and its binding partners are useful for medical imaging
of sites expressing (II). (I) and (II) are useful for treating disorders
involving aberrant protein expression or biological activity. The
polypeptide and polynucleotide sequences have applications in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits to assess biodiversity
and to produce other types of data and products dependent on DNA and
amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
coding sequences of the invention. Note: The sequence data for this
patent did not appear in the printed specification, but was obtained in
electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences
XX
```

XX	SQ	Sequence 6394 BP; 1840 A; 1597 C; 1384 G; 1571 T; 0 U; 2 Other;	
		Query Match 9.9%; Score 750; DB 5; Length 6394;	
		Best Local Similarity 99.8%; Pred. No. 0;	
		Matches 920; Conservative 0; Mismatches 1; Indels 1; Gaps 1;	
QY	1476	GAATTAATCAATGATGTCACCATACACAGGGGAGGAGAAATCTACTGCTTTC	1535
DB	3463	GAATTAATCAATGATGTCACCATACACAGGGGAGGAGAAATCTACTGCTTTC	3522
QY	1536	TGGAGAGACTAAGGGAGGCAITGAGGAGCGTGCCTCTCTGTCACCTGACTTCTCTGAAG	1595
DB	3523	TGGAGAGACTAAGGGAGGCAITGAGGAGCGTGCCTCTCTGTCACCTGACTTCTCTGAAG	3582
QY	1596	GCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAACT	1655
DB	3583	GCCAACTAATCTTAAAGCGTAAGTTTATCACTCAGTCAGCTGCAGACATTAGAAAAACT	3642
QY	1656	TCAAAAGTCTCGGTAGGCGCGGCAAACTTAGAAACCTATTGAACTTGGCACTC	1715
DB	3643	TCAAAAGTCTCGGTAGGCGCGGCAAACTTAGAAACCTATTGAACTTGGCACTC	3702
QY	1716	GGTTTTTATAATAGAGATCAGGAGGAGCAGCGGAAACAGGACAAACGGGATTTAAAAAA	1775
DB	3703	GGTTTTTATAATAGAGATCAGGAGGAGCAGCGGAAACAGGACAAACGGGATTTAAAAAA	3762
QY	1776	AGGCCAACCCCTTTAGTCATGACCCCTCAGGCAAGTGACCTTTGGAGGCTCGAAGGGA	1835
DB	3763	AGGCCAACCCCTTTAGTCATGACCCCTCAGGCAAGTGACCTTTGGAGGCTCGAAGGGA	3822
QY	1836	AAAGCTGGGCAATTTGAATGCTTAATAGGCTTGTTCAGTCCGCTCTAGAACGACT	1895
DB	3823	AAAGCTGGGCAATTTGAATGCTTAATAGGCTTGTTCAGTCCGCTCTAGAACGACT	3882
QY	1896	TTAAAAAGATTGTCCAACTAGAAGTAAGCCGCCCTTCGTCCATGCCCTTTATTTCAAG	1955
DB	3883	TTAAAAAGATTGTCCAACTAGAAGTAAGCCGCCCTTCGTCCATGCCCTTTATTTCAAG	3942
QY	1956	GGAATCACTGGAAGGCCACTGCCCGGAGCAAAAGGTCTTTTGTAGTCAGAAAGCACTA	2015
DB	3943	GGAATCACTGGAAGGCCACTGCCCGGAGCAAAAGGTCTTTTGTAGTCAGAAAGCACTA	4002
QY	2016	ACCAGATGATCAGCAGCAGGACTGAGGTGCTTGGGCAAGCGCCATCCATGCCATCA	2075
DB	4003	ACCAGATGATCAGCAGCAGGACTGAGGTGCTTGGGCAAGCGCCATCCATGCCATCA	4062
QY	2076	CCCTCACAGAGCCCTGGGTATGCTTGACCATTTGAGGGCCAGGAAGTGTCTCTCTGGACA	2135
DB	4063	CCCTCACAGAGCCCTGGGTATGCTTGACCATTTGAGGGCCAGGAAGTGTCTCTCTGGACA	4122
QY	2136	CTGGTGGGTCTTCTTGTAGTCTTCTTCTGTCCTGGGCAACTGCTCCAGATCTGTCA	2195
DB	4123	CTGGTGGGTCTTCTTGTAGTCTTCTTCTGTCCTGGGCAACTGCTCCAGATCTGTCA	4182
QY	2196	CTATTCTGAGGGGTCCTTAAGACGGGAGTCACTAGATACCTTTTCCAGGCACTAAGT	2255
DB	4183	CTATTCTGAGGGGTCCTTAAGACGGGAGTCACTAGATACCTTTTCCAGGCACTAAGT	4241
QY	2256	TATGAATCGGGAGCTTTATCTTTTTCATGCTTTTCTTAATATGCTTGAAGCCCCAC	2315
DB	4242	TATGAATCGGGAGCTTTATCTTTTTCATGCTTTTCTTAATATGCTTGAAGCCCCAC	4301
QY	2316	TACCTTTGTTAGGAGAGACATTTAGCAAAAGCAGGGGCCATTATACCTGAACATAGG	2375
DB	4302	TACCTTTGTTAGGAGAGACATTTAGCAAAAGCAGGGGCCATTATACCTGAACATAGG	4361
QY	2376	AGAAGGAACACCGTTTGTGT 2397	
DB	4362	AGAAGGAACACCGTTTGTGT 4383	

XX	DE	HERV R393E/Q envelope protein DNA SEQ ID NO 115.	
XX	KW	virucide; cytostatic; immunosuppressive; antiallergic; vaccine; screening; immune modulation; graft rejection; allergy; autoimmune disease; cancer; viral infection; envelope; env; gene; ds; mutant.	
XX	OS	Human endogenous retrovirus W.	
XX	PN	WO2005095442-A1.	
XX	PD	13-OCT-2005.	
XX	PF	30-MAR-2005; 2005WO-EP003339.	
XX	PR	30-MAR-2004; 2004EP-00290838.	
XX	PA	(INSR ) INST ROUSSY GUSTAVE.	
XX	PA	(CNRS ) CENT NAT RECH SCI.	
XX	PA	(UYPA-) UNIV PARIS SUD.	
XX	PI	Renard M, Mangeney M, Heidmann T;	
XX	XX	WPI; 2005-703311/72.	
XX	DR	P-PSDB; AED10985.	
XX	PT	New polypeptide capable of modulating the immunosuppressive properties of a viral protein or its fragment, useful for the manufacture of a vaccine for treating or preventing viral infections or cancer.	
XX	PS	Claim 41; SEQ ID NO 115; 262pp; English.	
CC	CC	The invention describes a polypeptide of 7-20 amino acid residues, which is capable of modulating the immunosuppressive properties of a viral protein or its fragment, against the host in which it is expressed (immunosuppression-modulatory sequence) when it substitutes the homologous sequence of the viral protein or its fragment, the polypeptide comprising the minimum following consensus amino acid sequence of (I).	
CC	CC	Also described are: a mutated ENV protein resulting from the mutation of a wild type ENV protein essentially carrying the following sequence: NY 1 Y 2 Y 3 LY 4 Y 5 LY 6 Y 7 Y 8 X 1 Y 9 Y 10 Y 11 CY 12 X 2 (II); a nucleic acid coding for the polypeptide or the mutated ENV protein; an eukaryotic or prokaryotic expression vector comprising the nucleic acid as well as the elements necessary for the expression of the nucleic acid; a recombinant cell comprising the nucleic acid or the eukaryotic or prokaryotic expression vector; a pharmaceutical or a vaccine composition comprising as active substance; at least one polypeptide, mutated ENV protein or its fragments, nucleic acid, prokaryotic or eukaryotic expression vector, or recombinant cell; in association with a pharmaceutically acceptable carrier; and antibodies or their fragments, scfv polypeptides, aptamers, or binding peptides, directed against mutated ENV proteins, provided that the antibodies or their fragments, scfv polypeptides, aptamers, or binding peptides do not bind to the corresponding wild type ENV proteins. The first mutation of a first amino acid and optionally of a second mutation of a second amino acid in a wild type viral envelope (ENV) protein is useful for manufacturing a mutated ENV protein having a modified immunosuppressive activity with respect to the wild type ENV protein. The protein is useful for the manufacture of a medicament or a vaccine intended for the prevention and/or the treatment of viral diseases, such as HTLV or FeLV infections; cancer; pathologies requiring an inhibition of the immune system, such as autoimmune diseases, allergies or graft rejections. The protein is also useful for screening compounds liable to modulate the immunosuppressive activity of viruses or tumor cells, where the compounds to screen are peptides, in particular peptides comprising from 5-30 amino acids, such as peptides originating from combinatorial peptide libraries. This sequence encodes	

6645 TGGGACATGGAAACGGGTGGCGACTCCCTGGTCACCTTGAAGATCAACTTAATCCCT 6704

CC	TYPE
CC	ENV

in having a modified immunosuppressive activity with respect

6645 TGGGACATGGAAACGGGTGGCGACTCCCTGGTCACCTTGAAGATCAACTTAATCCCT 6704

CC the wild type ENV protein. The protein is useful for the manufacture of a  
CC medicament or a vaccine intended for the prevention and/or the treatment  
CC of viral diseases, such as HTLV or FeLV infections; cancer; pathologies  
CC requiring an inhibition of the immune system, such as autoimmune  
CC diseases, allergies or graft rejections. The protein is also useful for  
CC screening compounds liable to modulate the immunosuppressive activity of  
CC viruses or tumor cells, where the compounds to screen are peptides, in  
CC particular peptides comprising from 5-30 amino acids, such as peptides  
CC originating from combinatorial peptide libraries. This sequence encodes  
CC HERV envelope protein.  
XX

SQ Sequence 1614 BP; 439 A; 452 C; 294 G; 423 T; 0 U; 6 Other;

Query Match 9.6%; Score 731; DB 14; Length 1614;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGAGAAATGACGGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 5744  
Db |||||  
QY 105 AGAGTTTCTATGAGAAATGACGGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCT 164  
Db |||||  
QY 5745 TTSTAAGGGAACCCCACTTCACTGCCCCACACCATATGCCCGCAACTGCTATCACTC 5804  
Db |||||  
QY 165 TTCTAAGGGAACCCCACTTCACTGCCCCACACCATATGCCCGCAACTGCTATCACTC 224  
Db |||||  
QY 5805 TGCCCACTCTTTCATGCATGCAAAATACCTATTATTGGACAGAAATAATGATTAACTCTAG 5864  
Db |||||  
QY 225 TGCCCACTCTTTCATGCATGCAAAATACCTATTATTGGACAGAAATAATGATTAACTCTAG 284  
Db |||||  
QY 5865 TTGTCTCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCCAACTGGTATGTC 5924  
Db |||||  
QY 285 TTGTCTCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCCAACTGGTATGTC 344  
Db |||||  
QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAATGTTAAAGAGTAACTCTCCA 5984  
Db |||||  
QY 345 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAATGTTAAAGAGTAACTCTCCA 404  
Db |||||  
QY 5985 ACTCACCGGGTACATGGGACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACA 6044  
Db |||||  
QY 405 ACTCACCGGGTACATGGGACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACA 464  
Db |||||  
QY 6045 TGAACACCTCCGTACCCATCTCGCTGGTAAAGCCTATTATATACACCCCTCACTGGGCT 6104  
Db |||||  
QY 465 TGAACACCTCCGTACCCATCTCGCTGGTAAAGCCTATTATATACACCCCTCACTGGGCT 524  
Db |||||  
QY 6105 CCATGAGGTCTCGGCCAAAAACCTACTAACTGTTGGATATGCCCTCCCGCTGAACTTCA 6164  
Db |||||  
QY 525 CCATGAGGTCTCGGCCAAAAACCTACTAACTGTTGGATATGCCCTCCCGCTGAACTTCA 584  
Db |||||  
QY 6165 GCATATGTTTCAATCCCTGTACCTGAAACAAATGGAACAACTTCAGCAGCAAAATAAACC 6224  
Db |||||  
QY 585 GCATATGTTTCAATCCCTGTACCTGAAACAAATGGAACAACTTCAGCAGCAAAATAAACC 644  
Db |||||  
QY 6225 CACTTCCGTTTATAGGACCTCTTGTTCCTCAATSTGGAATAACCCATACCTCAAACT 6284  
Db |||||  
QY 645 CACTTCCGTTTATAGGACCTCTTGTTCCTCAATCTGGAATAACCCATACCTCAAACT 704  
Db |||||  
QY 6285 CACTTGTGTAAATTTAGCAATFACATACACCAAACTCCCAATGCATCAGGTGGGT 6344  
Db |||||  
QY 705 CACTTGTGTAAATTTAGCAATFACATACACCAAACTCCCAATGCATCAGGTGGGT 764  
Db |||||  
QY 6345 AACTCTCCACACAAATAGTCTGCTACCTCAGCAATATTTTGTGTGGTACCTC 6404  
Db |||||  
QY 765 AACTCTCCACACAAATAGTCTGCTACCTCAGCAATATTTTGTGTGGTACCTC 824  
Db |||||  
QY 6405 AGCCTATCGTTGTTGAATGGCTCTTCAAACTATGCTCTCAATCTTATAGTGCC 6464  
Db |||||  
QY 825 AGCCTATCGTTGTTGAATGGCTCTTCAAACTATGCTCTCAATCTTATAGTGCC 884  
Db |||||  
QY 6465 CCCATATGCCATCTACCTAGAAACAAGATTATACAGTTATGATATCTTAAGCCCCGAA 6524  
Db |||||  
QY 885 CCCTATGACCATCTACCTAGAAACAAGATTATACAGTTATGATATCTTAAGCCCCGAA 944  
Db |||||

QY 6525 CAAAAGAGTACCCATCTTCTCTTTTATAGGACGAGGTGCTAGGTGCACCTAGGTAC 6584  
Db |||||  
QY 945 CAAAAGAGTACCCATCTTCTCTTTTATAGGACGAGGTGCTAGGTGCACCTAGGTAC 1004  
Db |||||  
QY 6585 TGGCATTTGGCGGTATACAACTCTCTACGTTCTACTACAACTATCTCAAGAACTAAA 6644  
Db |||||  
QY 1005 TGGCATTTGGCGGTATACAACTCTCTACGTTCTACTACAACTATCTCAAGAACTAAA 1064  
Db |||||  
QY 6645 TGGGAGCATGAAACGGGTCCGCGACTCCCTGGTCACTTTGCAAGATCAACTTAACCTCCT 6704  
Db |||||  
QY 1065 TGGGAGCATGAAACGGGTCCGCGACTCCCTGGTCACTTTGCAAGATCAACTTAACCTCCT 1124  
Db |||||  
QY 6705 AGCAGCAGTACTCTTTC 6721  
Db |||||  
QY 1125 AGCAGCAGTACTCTTTC 1141  
Db |||||

RESULT 14

AAH20070

ID AAH20070 standard; DNA; 1617 BP.

XX AAH20070;

XX 08-AUG-2001 (first entry)

XX HERV-W envelope protein G coding sequence.

XX Human endogenous retrovirus; HERV-W; HERV; chromosome 7; env protein;  
KW envelope protein; multiple sclerosis-related superantigen; vaccine;  
KW surface antigen; transmembrane; multiple sclerosis; neuroprotective;  
KW antisense-therapy; autoimmune disorder; ds.

XX Human endogenous retrovirus.

XX Key Location/Qualifiers

FD 1. .1617

FT /\*tag= a

FT /product= "HERV-W envelope protein G"

XX WO200131021-A1.

XX 03-MAY-2001.

XX 30-OCT-2000; 2000WO-EP010659.

XX 28-OCT-1999; 99EP-00402690.

XX (UYGE-) UNIV GENEVE.

XX Conrad B, Mach B;

XX WPI; 2001-316336/33.

XX P-PSDB; AAB75138.

XX New human retrovirus HERV-W ENV proteins/peptides having superantigen  
XX activity useful for diagnosing and treating multiple sclerosis.

XX Claim 13; Fig 10; 94pp; English.

XX On the basis of the PBS t-RNA motif used for the classification of human  
XX endogenous retrovirus (HERVs) the full length endogenous provirus which  
XX was located on the long arm of human chromosome 7 (7q21-22) has been  
XX designated HERV-W. The present invention describes proteins or peptides  
XX (I) having superantigen (SAG) activity comprising the ENV protein (ENV)  
XX of HERV-W, the surface protein (SU) and transmembrane (TM) sub-units. (I)  
XX have neuroprotective activity, and can be used in: vaccines; antisense-  
XX therapy; and HERV-W SAG activity-inhibitors. (I) and encoding DNA/RNA are  
XX useful for diagnosing multiple sclerosis (MS) or HERV-W-associated  
XX disorders. (I) are also useful for identifying substances (and optionally  
XX recovering) capable of binding to a retroviral superantigen associated  
XX with MS, substances capable of blocking SAG activity and substances  
XX capable of blocking transcription or translation of HERV-W retroviral  
XX superantigen. A protein or peptide derived from (I), modified to be

CC devoid of SAg activity and being capable of generating an immune response  
CC against HERV-W retroviral SAg is useful in therapy. Nucleic acid  
CC molecules encoding (I) are useful as vaccines against MS. Substances  
CC capable of blocking SAg activity, capable of binding to a retroviral  
CC superantigen associated with MS, or capable of blocking transcription or  
CC translation of HERV-W retroviral superantigen, for use in treating or  
CC preventing MS, obtained using (I) are useful for the treatment and  
CC prevention of MS. (I) and nucleic acids encoding them are useful for  
CC diagnosing autoimmune disease. The present sequence encodes the  
CC specifically claimed envelope protein of HERV-W designated G  
XX  
SQ Sequence 1617 BP; 442 A; 452 C; 296 G; 427 T; 0 U; 0 Other;

Query Match	9.6%	Score 731; DB 5; Length 1617;
Best Local Similarity	99.4%	Pred. No. 0;

5685	QY	AGAGTTTCTATGAGGAATGCAGCGTCGCCGAAATATTGTATGCCCATCGTATAGAGTCT	5744
105	DB	AGAGTTTCTATGAGGAATGCAGCGTCGCCGAAATATTGTATGCCCATCGTATAGAGTCT	164
5745	QY	TTSAAAGGGAACCCACCCTTCACTGCCCACACCATATGCCCGCAACTGCTATCACTC	5804
165	DB	TTCTAAGGGAAACCCACCCTTCACTGCCCACACCATATGCCCGCAACTGCTATCACTC	224
5805	QY	TGCCACTCTTTTGCAATGCATCAAATACTCAATTATTGGACAGGAAAAATGAATTAATCCTAG	5864
225	DB	TGCCACTCTTTTGCAATGCATCAAATACTCAATTATTGGACAGGAAAAATGAATTAATCCTAG	284
5865	QY	TTTCTCTGGAGGACTTGGAGTCACTGCTGTGGACHTTCACTTCAACCAAACTGGTATGTC	5924
285	DB	TTTCTCTGGAGGACTTGGAGTCACTGCTGTGGACHTTCACTTCAACCAAACTGGTATGTC	344
5925	QY	TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAACAATGTAAAGAAGTAATCTCCCA	5984
345	DB	TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAACAATGTAAAGAAGTAATCTCCCA	404
5985	QY	ACTCACSSGGTACATGGCACTTACGCCCTCAAAAGGACTAGATCTCTCAAAACTPACA	6044
405	DB	ACTCACSSGGTACATGGCACTTACGCCCTCAAAAGGACTAGATCTCTCAAAACTPACA	464
6045	QY	TGAAACCCCTCGTACCACACTCGCTGGTAAAGCCTATTTAATACCACCTCACTGGGCT	6104
465	DB	TGAAACCCCTCGTACCACACTCGCTGGTAAAGCCTATTTAATACCACCTCACTGGGCT	524
6105	QY	CCATGAGGCTCTGGCCCCAAACCCCTACTACTGTGGATATGCTCCCCCTGAACTTCAR	6164
525	DB	CCATGAGGCTCTGGCCCCAAACCCCTACTACTGTGGATATGCTCCCCCTGAACTTCAG	584
6165	QY	GCCATATGTTCAATCCCTGTACTGNAAGATGGAACTTCAAGCACAGAAATAAACAC	6224
585	DB	GCCATATGTTCAATCCCTGTACTGNAAGATGGAACTTCAAGCACAGAAATAAACAC	644
6225	QY	CAC TTCGGTTTTAGTAGGACCTCTGTGTTCCAAATSTGGAATAAACCATACCTCAACCT	6284
645	DB	CAC TTCGGTTTTAGTAGGACCTCTGTGTTCCAAATSTGGAATAAACCATACCTCAACCT	704
6285	QY	CACCTGTGTAATTTAGCAATACATACACAACAACCTCCCAATGCATCAGGTGGGT	6344
705	DB	CACCTGTGTAATTTAGCAATACATACACAACAACCTCCCAATGCATCAGGTGGGT	764
6345	QY	AAC TCCTCCACACAATAGTCGCTACCCCTCAGGAATATTTTGTCTGTGCTGCTACCTC	6404
765	DB	AAC TCCTCCACACAATAGTCGCTACCCCTCAGGAATATTTTGTCTGTGCTGCTACCTC	824
6405	QY	AGCCTATCGTGTGTTGAATGGCTCTTCAGAAATCTATGTGCTGCTCTCATCTTTAGTGCC	6464
825	DB	AGCCTATCGTGTGTTGAATGGCTCTTCAGAAATCTATGTGCTGCTCTCATCTTTAGTGCC	884
6465	QY	CCC YATGRCCACTCACTGAAACAAGATTTATPACAGTTATGTCATATCTTAAGCCCGCAA	6524
885	DB	CCC YATGRCCACTCACTGAAACAAGATTTATPACAGTTATGTCATATCTTAAGCCCGCAA	944

QY	6525	CAAAAGAGTACCCATTCTTCCTTTCTGTTATAGAGCAGAGTGTAGGTGCAC	TAGGTAC	6588
Db	945	CAAAAGAGTACCCATTCTTCCTTTCTGTTATAGAGCAGAGTGTAGGTGCAC	TAGGTAC	1004
QY	6585	TGGCATTGGCGGTATCACAACTCTACTCAGTCTTACTACAAACTATCTCAAGAACTAAA	6644	
Db	1005	TGGCATTGGCGGTATCACAACTCTACTCAGTCTTACTACAAACTATCTCAAGAACTAAA	1064	
QY	6645	TGGGGACATGGACGGGTGCCGACTCCCTGGTCACCTTGCAGATCAACTTAACTCCCT	6704	
Db	1065	TGGGGACATGGACGGGTGCCGACTCCCTGGTCACCTTGCAGATCAACTTAACTCCCT	1124	
QY	6705	AGCAGCAGTAGTCTCTTC	6721	
Db	1125	AGCAGCAGTAGTCTCTTC	1141	
RESULT 15				
AED10942				
ID	AED10942 standard; DNA; 1617 BP.			
XX				
AC	AED10942;			
XX				
XX	15-DEC-2005 (first entry)			
XX				
DE	HERV envelope protein DNA SEQ ID NO 73.			
XX				
KW	viricide; cytostatic; immunosuppressive; antiallergic; vaccine;			
KW	screening; immune modulation; graft rejection; allergy;			
KW	autoimmune disease; cancer; viral infection; envelope; env; gene; ds.			
XX				
OS	Human endogenous retrovirus W.			
XX				
PN	WO2005095442-A1.			
XX				
PD	13-OCT-2005.			
XX				
PF	30-MAR-2005; 2005WO-EP003339.			
XX				
PR	30-MAR-2004; 2004EP-00290838.			
XX				
PR	(INSR ) INST ROUSSY GUSTAVE.			
PA	(CNRS ) CENT NAT RECH SCI.			
PA	(IUPA-) UNIV PARIS SUD.			
XX				
PI	Renard M, Mangeney M, Heidmann T;			
XX				
XX	WPI: 2005-703311/72.			
DR	P-PSDB; AED10943.			
DR				
PT	New polypeptide capable of modulating the immunosuppressive properties of			
PT	a viral protein or its fragment, useful for the manufacture of a vaccine			
PT	for treating or preventing viral infections or cancer.			
XX				
PS	Disclosure; SEQ ID NO 73; 262pp; English.			
XX				
CC	The invention describes a polypeptide of 7-20 amino acid residues, which			
CC	is capable of modulating the immunosuppressive properties of a viral			
CC	protein or its fragment, against the host in which it is expressed			
CC	(immunosuppression-modulatory sequence) when it substitutes the			
CC	homologous sequence of the viral protein or its fragment, the polypeptide			
CC	comprising the minimum following consensus amino acid sequence of (I).			
CC	Also described are: a mutated ENV protein resulting from the mutation of			
CC	a wild type ENV protein essentially carrying the following sequence: NY 1			
CC	Y 2 Y 3 IY 4 Y 5 LY 6 Y 7 Y 8 X 1 Y 9 Y 10 Y 11 CY 12 X 2 (II); a nucleic			
CC	acid coding for the polypeptide or the mutated ENV protein; an eukaryotic			
CC	or prokaryotic expression vector comprising the nucleic acid as well as			
CC	the elements necessary for the expression of the nucleic acid; a			
CC	recombinant cell comprising the nucleic acid or the eukaryotic or			
CC	prokaryotic expression vector; a pharmaceutical or a vaccine composition			
CC	comprising as active substance: at least one polypeptide, mutated ENV			
CC	protein or its fragments, nucleic acid, prokaryotic or eukaryotic			
CC	expression vector, or recombinant cell; in association with a			





	Best Local Similarity	99.4%; Pred. No. 0;	Matches 1031; Conservative	0; Mismatches	6; Indels	0; Gaps	0;
Qy	5685	AGAGTTTCTATGCGAAGATGCAGCGTC	CCGGGAATAATTGATGCCCATCGTAGAGTGCT	5744			
Db	823	AGAGTTTCTATGCGAAGATGCAGCGTC	CCGGGAATAATTGATGCCCATCGTAGAGTGCT	882			
Qy	5745	TTSTAAGGAAACCCCAACTTTCATCTG	GCCCACACCACATATGCCCCGCAACTGCTATCACTC	5804			
Db	883	TTSTAAGGAAACCCCAACTTTCATCTG	GCCCACACCACATATGCCCCGCAACTGCTATCACTC	942			
Qy	5805	TGCCACTCTTTGCAATGCATGCAGAAAT	CTCATATTATGGACAGGAAAAATGATTAATCTCTAG	5864			
Db	943	TGCCACTCTTTGCAATGCATGCAGAAAT	CTCATATTATGGACAGGAAAAATGATTAATCTCTAG	1002			
Qy	5865	TTGTCTCTGGAGGACTTGGAGTCACTGT	CTGTGTGACCTTACTTCAACCCAACTGCTATGTC	5924			
Db	1003	TTGTCTCTGGAGGACTTGGAGTCACTGT	CTGTGTGACCTTACTTCAACCCAACTGCTATGTC	1062			
Qy	5925	TGATGGGGGTGGAGTTCAAGATCAGGCA	AGGAAAAAACATGTAAAAGAAAGTAATCTCCCA	5984			
Db	1063	TGATGGGGGTGGAGTTCAAGATCAGGCA	AGGAAAAAACATGTAAAAGAAAGTAATCTCCCA	1122			
Qy	5985	ACTCACCSGGGTACATGGCACTCTAGCC	CTACAAAGGACTAGATCTCTCAAACCTACA	6044			
Db	1123	ACTCACCSGGGTACATGGCACTCTAGCC	CTACAAAGGACTAGATCTCTCAAACCTACA	1182			
Qy	6045	TGAACCCCTCCGTACCCATYACTCGCTC	TGGTAAGGCTATTTAAATACCAACCCTCACTGGGCT	6104			
Db	1183	TGAACCCCTCCGTACCCATYACTCGCTC	TGGTAAGGCTATTTAAATACCAACCCTCACTGGGCT	1242			
Qy	6105	CCATGAGGTCTCGGCCCAAAACCCCTA	CTAACTGTTGGATATGCTCCCCCTGAACCTTCAR	6164			
Db	1243	CCATGAGGTCTCGGCCCAAAACCCCTA	CTAACTGTTGGATATGCTCCCCCTGAACCTTCAG	1302			
Qy	6165	GCCATATGTTTTCAATCCCTGACTGAA	CAATGGAACAACCTTGACGACAGAATAAACAC	6224			
Db	1303	GCCATATGTTTTCAATCCCTGACTGAA	CAATGGAACAACCTTGACGACAGAATAAACAC	1362			
Qy	6225	CAC TTCGGTTTTAGTAGGACCTCTGTG	TTCCTTCCCAATSTGGAATACCCCATACCTCCAACCT	6284			
Db	1363	CAC TTCGGTTTTAGTAGGACCTCTGTG	TTCCTTCCCAATSTGGAATACCCCATACCTCCAACCT	1422			
Qy	6285	CACCTGTGTAAAAATTTAGCAATACTA	CATACACAAACCACTCCCAATGCATCAGGTGGGT	6344			
Db	1423	CACCTGTGTAAAAATTTAGCAATACTA	CATACACAAACCACTCCCAATGCATCAGGTGGGT	1482			
Qy	6345	AACCTCTCCACACAAATAGTCTGCTAC	CCCTCAGGAATAATTTTGTCTGTGGTACCTC	6404			
Db	1483	AACCTCTCCACACAAATAGTCTGCTAC	CCCTCAGGAATAATTTTGTCTGTGGTACCTC	1542			
Qy	6405	AGCTATCGTTGTTTGAATGGCTCTTC	CAGAACTCTATGTGCTTCTCTCATCTCTTAGTGCC	6464			
Db	1543	AGCTATCGTTGTTTGAATGGCTCTTC	CAGAACTCTATGTGCTTCTCTCATCTCTTAGTGCC	1602			
Qy	6465	CCCYATGRCCATCTACACTGAACAAG	ATTTATACAGTTATGTGTATCTAATCTAAGCCCCGCAA	6524			
Db	1603	CCCYATGRCCATCTACACTGAACAAG	ATTTATACAGTTATGTGTATCTAATCTAAGCCCCGCAA	1662			
Qy	6525	CAAAAGAGTACCCANTTCTCTCTTTG	TATAGGACAGGAGTGCTAGGTGCATAGGTAC	6584			
Db	1663	CAAAAGAGTACCCANTTCTCTCTTTG	TATAGGACAGGAGTGCTAGGTGCATAGGTAC	1722			
Qy	6585	TGSCATTTGGCGGTATACACACCTCTA	CTCAGTTCTACTACAAATCTCTCAAGACTAAA	6644			
Db	1723	TGSCATTTGGCGGTATACACACCTCTA	CTCAGTTCTACTACAAATCTCTCAAGACTAAA	1782			
Qy	6645	TGGGGACATGGAAACGGGTCCGCGACT	CCCTGGGTCACTTCAAGATCAACTTAACTCCCT	6704			
Db	1783	TGGGGACATGGAAACGGGTCCGCGACT	CCCTGGGTCACTTCAAGATCAACTTAACTCCCT	1842			
Qy	6705	AGCAGCAGTAGTCTTCT	6721				

Db 1843 AGCAGCAGTAGTCTTC 1859

RESULT 17  
 AAAS59209

ID AAAS59209 standard; DNA; 1948 BP.

XX  
 AC AAAS59209;

XX  
 DT 07-NOV-2000 (first entry)

XX  
 DE 5' non coding, 3' pol gene, and 5' env gene of HERV-W from human genome.

XX  
 KW Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;

XX  
 KW gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.

XX  
 OS Homo sapiens.

XX  
 PN WO2000043521-A2.

XX  
 PD 27-JUL-2000.

XX  
 PF 21-JAN-2000; 2000WO-FR000144.

XX  
 PR 21-JAN-1999; 99FR-00000888.

XX  
 PA (INMR ) BIO MERIEUX.

XX  
 PI Paranhos-Baccala G, Mallet F, Voisset C;

XX  
 DR WPI; 2000-499229/44.

XX  
 PT New nucleic acid from human endogenous retrovirus, useful e.g. for

XX  
 PT diagnosis of autoimmune disease and complications of pregnancy, contains

XX  
 PT at least part of the gag gene.

XX  
 PS Disclosure; Page 45; 53pp; French.

XX  
 CC The present sequence represents an endogenous retroviral nucleic acid

XX  
 CC fragment, which is associated with an autoimmune disease, and is

XX  
 CC integrated into the human genome. The fragment is originally derived from

XX  
 CC a novel retrovirus, human endogenous retrovirus W (HERV-W). The HERV-W

XX  
 CC retrovirus is associated with autoimmune disease, failure of pregnancy or

XX  
 CC disorders of pregnancy. The nucleic acid fragment, or proteins derived

XX  
 CC from it, are useful for diagnosis of autoimmune disease (specifically

XX  
 CC multiple sclerosis) and for monitoring pregnancy. The nucleic acid

XX  
 CC fragments may also be used for in situ labelling of isolated chromosomes,

XX  
 CC while the transcription product can be used to study or monitor T cell

XX  
 CC proliferation in vitro

XX  
 SQ Sequence 1948 BP; 505 A; 549 C; 381 G; 510 T; 0 U; 3 Other;

Query Match 9.6%; Score 731; DB 3; Length 1948;  
 Best Local Similarity 99.4%; Pred. No. 0;  
 Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 5685 AGAGTTTCTATGAGAAATGACGCTCCCGGAAATATTGATGCCCATCGTATAGAGTCT 5744  
 |||||

Db 823 AGAGTTTCTATGAGAAATGACGCTCCCGGAAATATTGATGCCCATCGTATAGAGTCT 882  
 |||||

Qy 5745 TTSTAAGGGGAACCCCACTTTCAGTCCCAACCCATATGCCCGCAACTGCTATCACTC 5804  
 |||||

Db 883 TTGTAAGGGGAACCCCACTTTCAGTCCCAACCCATATGCCCGCAACTGCTATCACTC 942  
 |||||

Qy 5805 TGGCACTCTTTGATGATGCAATACTCAATTATGGACAGGAAAAATGATTAATCTAG 5864  
 |||||

Db 943 TGGCACTCTTTGATGATGCAATACTCAATTATGGACAGGAAAAATGATTAATCTAG 1002  
 |||||

Qy 5865 TTGTCTCGGAGGACTTGGAGTCACTGCTGTGGACTTACTTCCACCAAACTGGTAGTC 5924  
 |||||

Db 1003 TTGTCTCGGAGGACTTGGAGTCACTGCTGTGGACTTACTTCCCAAACTGGTAGTC 1062  
 |||||

Qy 5925 TGATGGGGTGGAGTTCCAAGATCAGGCAAGAGAAAAACATGTAAGAAAGTAATCTCCCA 5984  
 |||||

Db 1063 TGATGGGGTGGAGTTCAAGATCAGCAGAGAGAAAACATGTAAAGAGTAATCTCCA 1122  
QY 5985 ACTCACCGGGTACATGGCACCTCTAGCCCTACAAAGACTAGATCTCTCAAAACTACA 6044  
Db 1123 ACTCACCGGGTACATGGCACCTCTAGCCCTACAAAGACTAGATCTCTCAAAACTACA 1182  
QY 6045 TGAACCCCTCGGTACCCATCTCGCTCGGTGAAGCCCTATTAATACACCCCTCACTGGGCT 6104  
Db 1183 TGAACCCCTCGGTACCCATCTCGCTCGGTGAAGCCCTATTAATACACCCCTCACTGGGCT 1242  
QY 6105 CCATGAGGTCTCGGCCCAAAACCTCTACTAACTGTGGATATGCTCCCTCGAAGTTCAR 6164  
Db 1243 CCATGAGGTCTCGGCCCAAAACCTCTACTAACTGTGGATATGCTCCCTCGAAGTTCAR 1302  
QY 6165 GCCATATGTTCAATCCCTGTACCTGGAACAATGGAACAACCTTCAGCAGAGAAATAAACAC 6224  
Db 1303 GCCATATGTTTCAATCCCTGTACCTGGAACAATGGAACAACCTTCAGCAGAGAAATAAACAC 1362  
QY 6225 CACTTCCGTTTGTAGTAGGACCTCTTGTTCGAATSTGGAATAACCCATACCTCAAACT 6284  
Db 1363 CACTTCCGTTTGTAGTAGGACCTCTTGTTCGAATSTGGAATAACCCATACCTCAAACT 1422  
QY 6285 CACTGTGTAAATTTAGCAATCTACTACATACACAACTCCCAATGCATCAGGTGGGT 6344  
Db 1423 CACTGTGTAAATTTAGCAATCTACTACATACACAACTCCCAATGCATCAGGTGGGT 1482  
QY 6345 AACTCCTCCACACAAATAGTCTGCCTACCTCAGGAATATTTTGTGTGTGACCTC 6404  
Db 1483 AACTCCTCCACACAAATAGTCTGCCTACCTCAGGAATATTTTGTGTGTGACCTC 1542  
QY 6405 AGCCTATCGTTTGTGAATGGCTCTTCAGAACTATATGTCTCTCATTTCTTAGTGCC 6464  
Db 1543 AGCCTATCGTTTGTGAATGGCTCTTCAGAACTATATGTCTCTCATTTCTTAGTGCC 1602  
QY 6465 CCYATGRCATCTACCTGGAACAAGATTTATACAGTTATGTATATCTAAGCCCGCAA 6524  
Db 1603 CCTATGGCCATCTACCTGGAACAAGATTTATACAGTTATGTATATCTAAGCCCGCAA 1662  
QY 6525 CAAAAGAGTACCCTATCTCTCTTTTGTATAGGACGAGGTGCTAGGTGCACTAGGTAC 6584  
Db 1663 CAAAAGAGTACCCTATCTCTCTTTTGTATAGGACGAGGTGCTAGGTGCACTAGGTAC 1722  
QY 6585 TGGCATTTGGCGGTATACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 6644  
Db 1723 TGGCATTTGGCGGTATACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 1782  
QY 6645 TGGGACATGGAACGGGTGCGGACTCCCTGGTCACTTGGCAAGATCAACTTAACTCCCT 6704  
Db 1783 TGGGACATGGAACGGGTGCGGACTCCCTGGTCACTTGGCAAGATCAACTTAACTCCCT 1842  
QY 6705 AGCAGCAGTAGTCCTTC 6721  
Db 1843 AGCAGCAGTAGTCCTTC 1859

RESULT 18

ABN97948

ID ABN97948 standard; DNA; 2055 BP.

XX AC ABN97948;

XX DT 01-AUG-2002 (first entry)

XX DE Human retroviral HERV-7q env coding sequence.

XX KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;

XX KW multiple sclerosis; ds.

XX OS Human endogenous retrovirus.

XX PN WO9967395-A1.

XX XX

PD 29-DEC-1999.  
XX 23-JUN-1999; 99WO-FR001513.  
XX PR 23-JUN-1998; 98FR-00007920.  
XX PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
XX PI Alliel PM, Perin J, Rieger F;  
XX DR WPI; 2000-160587/14.  
XX PT New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used  
XX PT for diagnosis, treatment and prevention of autoimmune and neurological  
XX PT diseases.  
XX PS Claim 3; Page 142-145; 225pp; French.  
XX CC The present invention relates to new nucleic acid sequences of human  
XX CC endogenous retrovirus, HERV-7q, which is located on chromosome 7q.  
XX CC Regulatory elements associated with HERV-7q may alter expression of other  
XX CC genes (even remote genes) on the same chromosome, inducing immunological  
XX CC and/or neurological changes (which may be pathological or protective/  
XX CC curative). HERV-7q peptides can be used to improve efficiency of the  
XX CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding  
XX CC sequences can be used in immunogenic or vaccinating compositions, for  
XX CC protection against autoimmune diseases, particularly multiple sclerosis.  
XX CC The peptides may also be used (by sequence comparison) to detect/identify  
XX CC endogenous retroviruses that are abnormally expressed in cancer,  
XX CC neuropathologies or other autoimmune diseases. The present sequence was  
XX CC used to illustrate the invention  
XX SQ Sequence 2055 BP; 576 A; 574 C; 376 G; 529 T; 0 U; 0 Other;  
Query Match 9.6%; Score 731; DB 3; Length 2055;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 5685 AGAGTTTCTATGGAGATGACAGCGTCCCGAAATATTGATGCCCATCGTATAGGAGTCT 5744  
Db 495 AGAGTTTCTATGGAGATGACAGCGTCCCGAAATATTGATGCCCATCGTATAGGAGTCT 554  
QY 5745 TTSTAAGGGAACCCCACTTCACTGCCCCACACCCATATGCCGCAACTGCTATCCTC 5804  
Db 555 TTCTAAGGGAACCCCACTTCACTGCCCCACACCCATATGCCGCAACTGCTATCCTC 614  
QY 5805 TGCCACTCTTTGCAATGCAATATCTCATTTATGGACAGAAAAATGATTAATCCTAG 5864  
Db 615 TGCCACTCTTTGCAATGCAATATCTCATTTATGGACAGAAAAATGATTAATCCTAG 674  
QY 5865 TTGTCCTGGAGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCCAACTGGTATGTC 5924  
Db 675 TTGTCCTGGAGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCCAACTGGTATGTC 734  
QY 5925 TGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAGTAATCTCCCA 5984  
Db 735 TGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAGTAATCTCCCA 794  
QY 5985 ACTCACCGGGTACATGGCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACA 6044  
Db 795 ACTCACCGGGTACATGGCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACA 854  
QY 6045 TGAACCCCTCGGTACCCATCTCGCTGGTGAAGCCCTATTTAATACACCCCTCACTGGGCT 6104  
Db 855 TGAACCCCTCGGTACCCATCTCGCTGGTGAAGCCCTATTTAATACACCCCTCACTGGGCT 914  
QY 6105 CCATGAGGTCTCGGCCCAAAACCTCTACTAACTGTGGATATGCTCCCTCGAAGTTCAR 6164  
Db 915 CCATGAGGTCTCGGCCCAAAACCTCTACTAACTGTGGATATGCTCCCTCGAAGTTCAR 974  
QY 6165 GCCATATGTTTCAATCCCTGTACCTGGAACAATGGAACAACCTTCAGCAGAGAAATAAACAC 6224  
Db 975 GCCATATGTTTCAATCCCTGTACCTGGAACAATGGAACAACCTTCAGCAGAGAAATAAACAC 1034

QY 6225 CACTTCGGTTTATAGGACCTCTTGTGTTTCCAAATSTGGAATACCCATACCTCAAACT 6284  
 Db |||||  
 QY 1035 CACTTCGGTTTATAGGACCTCTTGTGTTTCCAAATCTGGAATAACCCATACCTCAAACT 1094  
 Db |||||  
 QY 6285 CACTGTGTAAATTTAGCAATACATACACAAACCACTCCCAATGCAATCAGGTGGT 6344  
 Db |||||  
 QY 1095 CACTGTGTAAATTTAGCAATACATACACAAACCACTCCCAATGCAATCAGGTGGT 1154  
 Db |||||  
 QY 6345 AACTCTCCACACAAATAGTGTGCTTACCTCAGGAATATTTTGTCTGTGTACCTC 6404  
 Db |||||  
 QY 1155 AACTCTCCACACAAATAGTGTGCTTACCTCAGGAATATTTTGTCTGTGTACCTC 1214  
 Db |||||  
 QY 6405 AGCTATCGTGTGTTGAATGCTCTCAGAACTCTATGTGCTTCTCTCATCTTAGTGC 6464  
 Db |||||  
 QY 1215 AGCTATCGTGTGTTGAATGCTCTCAGAACTCTATGTGCTTCTCTCATCTTAGTGC 1274  
 Db |||||  
 QY 6465 CCVATGRCATCTACACTGAAACAAGATTTATACAGTTATGTCTATCTAAGCCCGCAA 6524  
 Db |||||  
 QY 1275 CCCTATGACCATCTACACTGACAGATTTATACAGTTATGTCTATCTAAGCCCGCAA 1334  
 Db |||||  
 QY 6525 CAAAAGAGTACCACTCTCTCTTTGTATAGGACGAGGTGCTAGGTGCCTAGTAC 6584  
 Db |||||  
 QY 1335 CAAAAGAGTACCACTCTCTCTTTGTATAGGACGAGGTGCTAGGTGCCTAGTAC 1394  
 Db |||||  
 QY 6585 TGGCATTGGCGGTATCAACACTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA 6644  
 Db |||||  
 QY 1395 TGGCATTGGCGGTATCAACACTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA 1454  
 Db |||||  
 QY 6645 TGGGGACATGGAACGGGTGCGGACTCCCTGCTGCTCAGCTTCAAGATCAACTTAACTCCCT 6704  
 Db |||||  
 QY 1455 TGGGGACATGGAACGGGTGCGGACTCCCTGCTGCTCAGCTTCAAGATCAACTTAACTCCCT 1514  
 Db |||||  
 QY 6705 AGCAGCAGTAGTCTTCT 6721  
 Db |||||  
 QY 1515 AGCAGCAGTAGTCTTCT 1531  
 Db |||||

## RESULT 19

AA25661

ID AAX25661 standard; cDNA to mRNA; 2782 BP.

AC AAX25661;

XX AAX25661;

XX 21-MAY-1999 (first entry)

XX Human endogenous retrovirus W clone cl.PH74.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

This sequence represents clone cl.PH74 of the human endogenous retrovirus (HERV) W genome. The nucleic acids, their fragments or peptides encoded by them are markers of autoimmune disease (e.g. multiple sclerosis, rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-dependent diabetes and related pathologies) and of abnormal or unsuccessful pregnancy and can be used as chromosomal markers for susceptibility to these conditions, or proximity markers of genes associated with this susceptibility

SQ Sequence 2782 BP; 741 A; 767 C; 565 G; 709 T; 0 U; 0 Other;

Query Match 9.6%; Score 731; DB 2; Length 2782;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 1031; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGAGAAATGACGGTCCCGGAAATATTTGATGCCCATCGTATAGGAGTCT 5744

Db |||||

QY 867 AGAGTTTCTATGAGAAATGACGGTCCCGGAAATATTTGATGCCCATCGTATAGGAGTCT 926

Db |||||

QY 5745 TTSTAAGGGAACCCCCACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 5804

Db |||||

QY 927 TTCTAAGGGAACCCCCACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 986

Db |||||

QY 5805 TGGCACTCTTGCATGCATGCAAAATCACTATTATGGACAGGAAATGATTAATCTCTAG 5864

Db |||||

QY 987 TGGCACTCTTGCATGCATGCAAAATCACTATTATGGACAGGAAATGATTAATCTCTAG 1046

Db |||||

QY 5865 TTGTCTCTGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCCACCCAACTGGTATGTC 5924

Db |||||

QY 1047 TTGTCTCTGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCCACCCAACTGGTATGTC 1106

Db |||||

QY 5925 TGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAGAAAAACATGTAAAAAGAGTAATCTCCCA 5984

Db |||||

QY 1107 TGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAGAAAAACATGTAAAAAGAGTAATCTCCCA 1166

Db |||||

QY 5985 ACTCACCSGGGTACATGGCACTCTAGCCCTCAAAAGAGACTAGATCTCTCAAAACTACA 6044

Db |||||

QY 1167 ACTCACCSGGGTACATGGCACTCTAGCCCTCAAAAGAGACTAGATCTCTCAAAACTACA 1226

Db |||||

QY 6045 TGAACCCCTCCGTACCCATCTCGCTGCTAGGCTTATTAATACCACTCTCACTGGCT 6104

Db |||||

QY 1227 TGAACCCCTCCGTACCCATCTCGCTGCTAGGCTTATTAATACCACTCTCACTGGCT 1286

Db |||||

QY 6105 CCATGAGGTCTCGGCCCAAAACCTACTAACTGTGGATATGCTCCCTCGTAACCTTCAR 6164

Db |||||

QY 1287 CCATGAGGTCTCGGCCCAAAACCTACTAACTGTGGATATGCTCCCTCGTAACCTTCAG 1346

Db |||||

QY 6165 GCATATGTTTCAATCCCTGTACCTGAAACAATGGAAACAATTCAGACAGAAATAAACAC 6224

Db |||||

QY 1347 GCATATGTTTCAATCCCTGTACCTGAAACAATGGAAACAATTCAGACAGAAATAAACAC 1406

Db |||||

QY 6225 CACTTCCGTTTATGAGGACCTCTTGTTCCTATSTGGAATACCCATACCTCAAACT 6284

Db |||||

QY 1407 CACTTCCGTTTATGAGGACCTCTTGTTCCTATSTGGAATACCCATACCTCAAACT 1466

Db |||||

QY 6285 CACTGTGTAAAAATTTAGCAATACATACACAACAACCACTCCCAATGATCAGGTGGT 6344

Db |||||

QY 1467 CACTGTGTAAAAATTTAGCAATACATACACAACAACCACTCCCAATGATCAGGTGGT 1526

Db |||||

QY 6345 AACTCCTCCACACAATAATAGTCTGCTACCCCTCAGGAATATTTTGTCTGTGGTACCTC 6404

Db |||||

QY 1527 AACTCCTCCACACAATAATAGTCTGCTACCCCTCAGGAATATTTTGTCTGTGGTACCTC 1586

Db |||||

QY 6405 AGCCTATCGTTTGTGAATGGCTCTTCAGAACTATGTGCTCTCTCATCTTCTTAGTGCC 6464

Db |||||

QY 1587 AGCCTATCGTTTGTGAATGGCTCTTCAGAACTATGTGCTCTCTCATCTTCTTAGTGCC 1646

Db |||||

QY 6465 CCCTATGRCATCTACCTGAAACAAGATTTATACAGTTATGTGCATATCTTAAGCCCGCAA 6524

Db |||||

QY 1647 CCCTATGRCATCTACCTGAAACAAGATTTATACAGTTATGTGCATATCTTAAGCCCGCAA 1706

Db |||||

QY 6525 CAAAAGAGTACCCATTTCTTCTTTTATAGGAGCAGGAGTGTAGGTGCATAGGTAC 6584

Db |||||

QY

QY

QY 6225 CACTTCGGTTTATAGGACCTCTTGTGTTTCCAAATSTGGAATACCCATACCTCAAACT 6284

Db |||||

QY 1035 CACTTCGGTTTATAGGACCTCTTGTGTTTCCAAATCTGGAATAACCCATACCTCAAACT 1094

Db |||||

QY 6285 CACTGTGTAAATTTAGCAATACATACACAAACCACTCCCAATGCAATCAGGTGGT 6344

Db |||||

QY 1095 CACTGTGTAAATTTAGCAATACATACACAAACCACTCCCAATGCAATCAGGTGGT 1154

Db |||||

QY 6345 AACTCTCCACACAAATAGTGTGCTTACCTCAGGAATATTTTGTCTGTGTACCTC 6404

Db |||||

QY 1155 AACTCTCCACACAAATAGTGTGCTTACCTCAGGAATATTTTGTCTGTGTACCTC 1214

Db |||||

QY 6405 AGCTATCGTGTGTTGAATGCTCTCAGAACTCTATGTGCTTCTCTCATCTTAGTGC 6464

Db |||||

QY 1215 AGCTATCGTGTGTTGAATGCTCTCAGAACTCTATGTGCTTCTCTCATCTTAGTGC 1274

Db |||||

QY 6465 CCVATGRCATCTACACTGAAACAAGATTTATACAGTTATGTCTATCTAAGCCCGCAA 6524

Db |||||

QY 1275 CCCTATGACCATCTACACTGACAGATTTATACAGTTATGTCTATCTAAGCCCGCAA 1334

Db |||||

QY 6525 CAAAAGAGTACCACTCTCTCTTTGTATAGGACGAGGTGCTAGGTGCCTAGTAC 6584

Db |||||

QY 1335 CAAAAGAGTACCACTCTCTCTTTGTATAGGACGAGGTGCTAGGTGCCTAGTAC 1394

Db |||||

QY 6585 TGGCATTGGCGGTATCAACACTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA 6644

Db |||||

QY 1395 TGGCATTGGCGGTATCAACACTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA 1454

Db |||||

QY 6645 TGGGGACATGGAACGGGTGCGGACTCCCTGCTGCTCAGCTTCAAGATCAACTTAACTCCCT 6704

Db |||||

QY 1455 TGGGGACATGGAACGGGTGCGGACTCCCTGCTGCTCAGCTTCAAGATCAACTTAACTCCCT 1514

Db |||||

QY 6705 AGCAGCAGTAGTCTTCT 6721

Db |||||

QY 1515 AGCAGCAGTAGTCTTCT 1531

Db |||||

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

QY

Db	1707	CAAAAGAGTACCCATTCTTCCTTTGTTATAGGACGAGGAGTGCTAGGTGCACCTAGGTAC	1766
Qy	6585	TGGGCAATTGGCGGGTATCACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAACTAAA	6644
Db	1767	TGGCAATTGGCGGGTATCACAACCTCTACTCTCAGTTCTACTACAAACTATCTCAAGAACTAAA	1826
Qy	6645	TGGGGACATGGAAACGGGTGCGCGGACTCGCTGGTCACTTGCAAGATCAACTTAACCTCCCT	6704
Db	1827	TGGGGACATGGAAACGGGTGCGCGGACTCGCTGGTCACTTGCAAGATCAACTTAACCTCCCT	1886
Qy	6705	AGCAGCAGTAGTCCTTC	6721
Db	1887	AGCAGCAGTAGTCCTTC	1903

RESULT 20  
AAA59211  
ID AAA59211 standard; DNA; 2782 BP.

AC AC  
XX AA59211;  
DT 07-NOV-2000 (first entry)

5' non coding, 3' pol gene, and 3' non coding sequences of HERV-W.  
XX  
KW Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;  
KW gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.

OS Homo sapiens.  
XX  
PN WC200043521-A2.

PD	27-JUL-2000.
XX	
PF	21-JAN-2000; 200WO-FR000144.

PR 21-JAN-1999; 99PR-00000888.  
XX  
PA (INMR ) BIO MERIEUX.

PI Paranhos-Baccala G, Mallet F, Voisset C;  
XX  
DR WPI; 2000-492229/44.

PT New nucleic acid from human endogenous retrovirus, useful e.g. for  
PT diagnosis of autoimmune disease and complications of pregnancy, contains  
PT at least part of the gag gene.

PS Disclosure; Page 46-47; 53pp; French.  
XX The present sequence represents an endogenetic retroviral nucleic acid

integrated into the human genome. The fragment is originally derived from a novel retrovirus, human endogenous retrovirus W (HERV-W). The HERV-W retrovirus is associated with autoimmune disease, failure of pregnancy or disorders of pregnancy. The nucleic acid fragment, or proteing derived from it, are useful for diagnosis of autoimmune disease (specifically multiple sclerosis) and for monitoring pregnancy. The nucleic acid fragments may also be used for *in situ* labelling of isolated chromosomes, while the transcription product can be used to study or monitor T cell proliferation *in vitro*.

SQ Sequence 2782 BP; 741 A; 767 C; 565 G; 709 T; 0 U; 0 Other;  
 Query Match 9.6%; Score 731; DB 3; Length 2782;  
 Best Local Similarity 99.4%; Pred. No. 0;  
 Matches 1031; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

<b>QY</b>	5685	AGAGTTTCTATGGGAATCAGCGTCCCGAAATATTGATGCCCATCGTATAGGAGTCT	5744
<b>Db</b>	867	AGAGTTTCTATGGGAATCAGCGTCCCGAAATATTGATGCCCATCGTATAGGAGTCT	926

Qy	5745	TTTSTAAGGGAAACCCCAACCTTCTCACTGCCACACCCCATATGCCCCCGAAATGCTGTATCACTC	5804
Db	927	TTCTAAAGGGAAACCCCAACCTTCTCACTGCCACACCCCATATGCCCCCGAAATGCTGTATCACTC	986
Qy	5805	TGCACCTCTTTGCATGCATGCACAAATACTCATTTATTCGACAGGAAAAATGATTAACTCTAG	5864
Db	987	TGCCACTCTTTGCATGCATGCACAAATACTCATTTATTCGACAGGAAAAATGATTAACTCTAG	1046
Qy	5865	TTGTCTCTGGAGGACTTTGGAGTCACTGTCTGTGTGGACTTACTTCCACCCAAACTGGTATGTC	5924
Db	1047	TTGTCTCTGGAGGACTTTGGAGTCACTGTCTGTGTGGACTTACTTCCACCCAAACTGGTATGTC	1106
Qy	5925	TGATGGGGGTGGAGTTCACAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCA	5984
Db	1107	TGATGGGGGTGGAGTTCACAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCA	1166
Qy	5985	ACTCACCSGGTACATGGCAACCTCTAGCCCTACAAAGGACTAGATCTCTCAAAACTACA	6044
Db	1167	ACTCACCGGGGTACATGGCAACCTCTAGCCCTACAAAGGACTAGATCTCTCAAAACTACA	1226
Qy	6045	TGAACCCCTCCGTAACCCATACTCGCTGGTGAAGCTATTATTAATACCAACCTCACTGGGCT	6104
Db	1227	TGAACCCCTCCGTAACCCATACTCGCTGGTGAAGCTATTATTAATACCAACCTCACTGGGCT	1286
Qy	6105	CCATGAGGTCTCGGCCCCAAAACCTACTACTCTGTTGGATATGCCTCCCCCTGAACCTCAR	6164
Db	1287	CCATGAGGTCTCGGCCCCAAAACCTACTACTCTGTTGGATATGCCTCCCCCTGAACCTCAG	1346
Qy	6165	GCCATATGTTTCAATCCCTGTACTCGCAACAATGGAAACAACTTCAGCACAGAAATAAACAC	6224
Db	1347	GCCATATGTTTCAATCCCTGTACTCGAACHAATGGAAACAACTTCAGCACAGAAATAAACAC	1406
Qy	6225	CACCTCGTTTTATGATAGGACCTCTGTGTTCCAAATSTGSAATAAACCATACCTCAAACCT	6284
Db	1407	CACCTCGTTTTATGATAGGACCTCTGTGTTCCAAATSTGSAATAAACCATACCTCAAACCT	1466
Qy	6285	CACCTGTGTAAATTTTAGCAATACTACATACACAAACCACTCCCAATGCAATCAGGTGGGT	6344
Db	1467	CACCTGTGTAAATTTTAGCANATCTACATACACAAACCACTCCCAATGCAATCAGGTGGGT	1526
Qy	6345	AACTCCTCCCAACAATAATGTCGCCTACCCCTCAGGAATATTTTTGTCTGTGGTACCTC	6404
Db	1527	AACTCCTCCCAACAATAATGTCGCCTACCCCTCAGGAATATTTTTGTCTGTGGTACCTC	1586
Qy	6405	AGCCTATCGTTGTTTGAATGGCTCTTCAGAACTCTATGTGCTTCTCTCATCTCTTAGTGCC	6464
Db	1587	AGCCTATCGTTGTTTGAATGGCTCTTCAGAACTCTATGTGCTTCTCTCATCTCTTAGTGCC	1646
Qy	6465	CCCVATGRCATCTACACTCAACAAGANTTTATACAGTTATGTGCATATCTAAGCCCCCGCAA	6524
Db	1647	CCCTATGACCATCTACACTGACAGANTTTATACAGTTATGTGCATATCTAAGCCCCCGCAA	1706
Qy	6525	CAAAAGAGTACCCATCTTCCTTTTGTATAGAGCAGGAGTGTAGGTGCACCTAGGTAC	6584
Db	1707	CAAAAGAGTACCCATCTTCCTTTTGTATAGAGCAGGAGTGTAGGTGCACCTAGGTAC	1766
Qy	6585	TGGCATTTGGGGGTATCAACACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA	6644
Db	1767	TGGCATTTGGGGGTATCAACACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA	1826
Qy	6645	TGGGGACATCGAAACGGGTCCCGCACTCCCTGGGTCACTTCGCAAGATCAAACTTAACTCCCT	6704
Db	1827	TGGGGACATCGAAACGGGTCCCGCACTCCCTGGGTCACTTCGCAAGATCAAACTTAACTCCCT	1886
Qy	6705	AGCAGCAGTAGTCCCTTC	6721
Db	1887	AGCAGCAGTAGTCCCTTC	1903

RESULT 21  
AAH20069  
ID AAH2  
XX



QY 6585 TGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 6644  
DB 1767 TGGCATTGGCGGTATCACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAA 1826  
QY 6645 TGGGACATGGAACGGGTGCGCGACTCCCTGTGTACCTTTGCAAGATCAACTTAACTCCCT 6704  
DB 1827 TGGGACATGGAACGGGTGCGCGACTCCCTGTGTACCTTTGCAAGATCAACTTAACTCCCT 1886  
QY 6705 AGCAGAGTAGTCCTTC 6721  
DB 1887 AGCAGAGTAGTCCTTC 1903

RESULT 22  
AAD24195  
ID AAD24195 standard; cDNA; 2930 BP.

AC AAD24195;  
DT 07-MAY-2002 (first entry)  
DE Human syncytin cDNA.  
KW Human; syncytin; preeclampsia; gestational trophoblast disorder;  
KW choriocarcinoma; hydatiform mole; placental site tumour; abortion;  
KW envelope gene; human endogenous defective retrovirus; HERV-W; ss.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 930..2546  
FT /\*tag= a  
FT /product= "syncytin"

PN WO200204678-A2.  
PD 17-JAN-2002.  
PF 09-JUL-2001; 2001WO-US021719.  
PR 07-JUL-2000; 2000US-0216657P.  
PA (GEMV ) GENETICS INST INC.  
PI Keith JC, McCoy JM, Mi S;  
DR WPI; 2002-171727/22.  
XX P-PSDB; AAE14540.  
PT Identifying a compound for treating a subject with or at risk of  
PT developing preeclampsia, comprises determining whether the expression or  
PT activity of syncytin in the cell is modulated in the presence of a test  
PT compound.  
XX Disclosure; Page 39-42; 43pp; English.

XX The invention relates to identifying compounds which are modulators of  
CC syncytin expression. The syncytin modulators are useful in diagnosis and  
CC treatment of preeclampsia and gestational trophoblast disorders (e.g.  
CC choriocarcinoma, hydatiform mole, placental site tumour and missed/  
CC incomplete abortion). Syncytin is a human gene derived from the envelope  
CC gene of human endogenous defective retrovirus, HERV-W. The present  
CC invention is based partly on the discovery that syncytin expression is  
CC dramatically reduced in preeclampsia, and is also mis-localised to the  
CC apical syncytiotrophoblast membrane. The present sequence is human  
CC syncytin cDNA

SQ Sequence 2930 BP; 842 A; 800 C; 571 G; 717 T; 0 U; 0 Other;  
Query Match 9.6%; Score 728; DB 6; Length 2930;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1898; Conservative 0; Mismatches 21; Indels 1; Gaps 1;

QY 4802 AGGCTTCCCTGTAGGACACAAAAGGCCCAAGAGTAAATAAAGSCACTAGTTTCATGAAA 4861  
DB 152 AGGCTTCCCTGTAGGACACAAAAGGCCCAAGAGTAAATAAAGSCACTAGTTTCATGAAA 211  
QY 4862 TAAATCCAGATTCCGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGG 4921  
DB 212 TAAATCCAGATTCCGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGG 271  
QY 4922 CCACAGTAAACCCAGGAGTATCCACGGGCTTAGGTATAGTATACATTCACATCGGCT 4981  
DB 272 CCACAGTAAACCCAGGAGTATCCACGGGCTTAGGTATAGTATACATTCACATCGGCT 331  
QY 4982 GAAGGCCACAGTCTCAGGGAAGTTCGAGAAATGAATGAAYACTCAAGAGACATCTAA 5041  
DB 332 GAAGGCCACAGTCTCAGGGAAGTTCGAGAAATGAATGAAYACTCAAGAGACATCTAA 391  
QY 5042 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCTGTCTGTGGCTATAGCTTAAAAA 5101  
DB 392 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCTGTCTGTGGCTATAGCTTAAAAA 451  
QY 5102 GAATCTGCAACTTTCCTCCAAAAGAGGACTTTAGCCCATAGAAATGCTGTATGAAGGC 5161  
DB 452 GAATCTGCAACTTTCCTCCAAAAGAGGACTTTAGCCCATAGAAATGCTGTATGAAGGC 511  
QY 5162 CTTTCATAAACAATGACTTGTGTGACCCCAAGACAGCAACTTAGTTGACAGACATCAC 5221  
DB 512 CTTTCATAAACAATGACTTGTGTGACCCCAAGACAGCAACTTAGTTGACAGACATCAC 571  
QY 5222 CTCTTTAGCCAAATCAACAAGTCTTTAAACATTTACAAGGAACCTATCCCTGAGAAGA 5281  
DB 572 CTCTTTAGCCAAATCAACAAGTCTTTAAACATTTACAAGGAACCTATCCCTGAGAAGA 631  
QY 5282 GGGAAAAAGAACTATTCCACCCCTTGTGACATGTTAGTCAAGTCCCTTCTCTAAATTC 5341  
DB 632 GGGAAAAAGAACTATTCCACCCCTTGTGACATGTTAGTCAAGTCCCTTCTCTAAATTC 691  
QY 5342 CCCATCCCTAGATACATCTCGGAGGACCTACCCAGTCAATTTTATTTACCCCAACTGC 5401  
DB 692 CCCATCCCTAGATACATCTCGGAGGACCTTACCCAGTCAATTTTATCTACCCCAACTGC 751  
QY 5402 GGTAAAGTGGCTGGAGTCTTGGATACATCACACTTGGAGTCAAAATCTCGGATCT 5461  
DB 752 GGTAAAGTGGCTGGAGTCTTGGATACATCACACTTGGAGTCAAAATCTCGGATCT 811  
QY 5462 GCCAAAGGAACTGAAATCCAGGAGACAAAGCTAGCTATTCTCTGGAACCTCTAGAGA 5521  
DB 812 GCCAAAGGAACTGAAATCCAGGAGACAAAGCTAGCTATTCTCTGGAACCTCTAGAGA 871  
QY 5522 TTTGGCGCTGCTCTTCAACACACACAGGAGGAAGTAACTAAATCATTAATCCCCA 5581  
DB 872 TTTGGCGCTGCTCTTCAACACACACAGGAGGAAGTAACTAAATCATTAATCCCCA 930  
QY 5582 TGGGCTCCCTTATCATATTTTCTKTASTGTSTTTTACCTSTTTTCACTCTCACATG 5641  
DB 931 TGGGCTCCCTTATCATATTTTCTKTASTGTSTTTTACCTSTTTTCACTCTCACATG 990  
QY 5642 CACCCCTCCATGCGCTGTATGACAGTAGTCCCTTACCMAGAGTTTCTATGAGAA 5701  
DB 991 CACCCCTCCATGCGCTGTATGACAGTAGTCCCTTACCMAGAGTTTCTATGAGAA 1050  
QY 5702 TGCAGGCTCCGGAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGGAACCCCCA 5761  
DB 1051 TGCAGGCTCCGGAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGGAACCCCCA 1110  
QY 5762 CTTTCACATGCCACACACCATATGCCCCGCAACTGTATCACCTCTGCACTCTTTGATGC 5821  
DB 1111 CTTTCACATGCCACACACCATATGCCCCGCAACTGTATCACCTCTGCACTCTTTGATGC 1170  
QY 5822 ATGCAAAATCACTATTATTGGACAGGAAAAATGAATTAATCTAGTTGCTCTGGAGGACTTG 5881  
DB 1171 ATGCAAAATCACTATTATTGGACAGGAAAAATGAATTAATCTAGTTGCTCTGGAGGACTTG 1230





Db	1423	ACCAGATGATCCAGCAGCAGCAGCATGAGGGTGCCCTGGGGCAAGCCCAATCCCATGCCCATCA	1488			
Qy	2076	CCCTCAGAGAGCCCTGGGTATGCTTTGACCACTTGAGGGCCAGGAAGGTGTCTCTCTGGACA	2135			
Db	1483	CCCTCAGAGAGCCCTGGGTATGCTTTGACCACTTGAGGGCCAGGAAGGTGTCTCTCTGGACA	1542			
Qy	2136	CTGTGCGGGTCTTCTTAACTCTTACTCTTCTGTCCCGGACAACTGTCTCCAGATCTGTCA	2195			
Db	1543	CTGTGCGGGTCTTCTTAACTCTTACTCTTCTGTCCCGGACAACTGTCTCCAGATCTGTCA	1602			
Qy	2196	CTAT 2199				
Db	1603	CTAT 1606				
RESULT 24						
AAAS9206						
ID	AAAS9206 standard; DNA; 2938 BP.					
XX	AAAS9206;					
XX	07-NOV-2000 (first entry)					
DT	Gag and partial pol gene fragment of HERV-W from human genome.					
DE	Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;					
KW	gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.					
KW	Homo sapiens.					
OS	WO200043521-A2.					
XX	27-JUL-2000.					
XX	21-JAN-2000; 2000WO-FR000144.					
PF	21-JAN-1999; 99FR-00000888.					
PR	(INMR ) BIO MERIEUX.					
XX	Paranhos-Baccala G, Mallet F, Voisset C;					
PA	WPI; 2000-499229/44.					
XX	New nucleic acid from human endogenous retrovirus, useful e.g. for					
PT	diagnosis of autoimmune disease and complications of pregnancy, contains					
PT	at least part of the gag gene.					
XX	Disclosure; Page 43; 53pp; French.					
PS	The present sequence represents an endogenous retroviral nucleic acid					
XX	fragment, which is associated with an autoimmune disease, and is					
CC	integrated into the human genome. The fragment is originally derived from					
CC	a novel retrovirus, human endogenous retrovirus W (HERV-W). The HERV-W					
CC	retrovirus is associated with autoimmune disease, failure of pregnancy or					
CC	disorders of pregnancy. The nucleic acid fragment, or proteins derived					
CC	from it, are useful for diagnosis of autoimmune disease (specifically					
CC	multiple sclerosis) and for monitoring pregnancy. The nucleic acid					
CC	fragments may also be used for in situ labelling of isolated chromosomes,					
CC	while the transcription product can be used to study or monitor T cell					
CC	proliferation in vitro					
XX	Sequence 2938 BP; 878 A; 720 C; 646 G; 692 T; 0 U; 2 Other;					
SQ	Query Match 9.5%; Score 724; DB 3; Length 2938;					
Best Local Similarity 100.0%; Pred. No. 0;						
Matches 724; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
Qy	1476	GAATTACTCAATGATGTCCACCATTAACACAGGGGAGGAGAAATCCTACTGCCTTTC	1535			
Db	883	GAATTACTCAATGATGTCCACCATTAACACAGGGGAGGAGAAATCCTACTGCCTTTC	942			
Qy	1536	TGGAGAGACTAAGGGAGGCAATTGAGGAAGCGTGCCTCTCTGTCACTGACTTCTTGAAG	1595			



CC ameliorating medical conditions in humans and animals, although no  
 CC supporting data is given. Suggested activities include nutritional  
 CC activity, cytokine and cell proliferation/differentiation activity,  
 CC immune stimulating (e.g. as vaccines) or suppressing activity,  
 CC haematopoiesis regulating activity, tissue growth activity,  
 CC activin/inhibin activity, chemotactic/chemokinetic activity,  
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 CC activity, cadherin/tumour invasion suppressor activity, and tumour  
 CC inhibition activity. The polynucleotides are also stated to be useful for  
 CC gene therapy  
 XX  
 SQ Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 U; 0 Other;  
 Query Match 9.0%; Score 680; DB 2; Length 2946;  
 Best Local Similarity 99.3%; Pred. No. 0;  
 Matches 1030; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 5685 AGAGTTTCTATGAGAAATGACGCTCCGGAATATTGATGCCCATCGTATAGGAGTCT 5744  
 DB 1032 AGAGTTTCTATGAGAAATGACGCTCCGGAATATTGATGCCCATCGTATAGGAGTCT 1091  
 QY 5745 TTSTAAGGGAACCCCACTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 5804  
 DB 1092 TTCTAAGGGAACCCCACTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTC 1151  
 QY 5805 TGCCACTCTTTCATGCATGCAAACTACTATTTGGACAGAAATGATTAATCCTAG 5864  
 DB 1152 TGCCACTCTTTCATGCATGCAAACTACTATTTGGACAGAAATGATTAATCCTAG 1211  
 QY 5865 TTGTCTCGAGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCAACTGGTATGTC 5924  
 DB 1212 TTGTCTCGAGACTTGGAGTCACTGTCTGTGGACTTACTTCAACCAACTGGTATGTC 1271  
 QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAAAGTAATCTCCA 5984  
 DB 1272 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAGAAAGTAATCTCCA 1331  
 QY 5985 ACTACCCGGGTACATGGCACTCTAGCCCTTACAAAGACTAGATCTCTCAAACTACA 6044  
 DB 1332 ACTACCCGGGTACATGGCACTCTAGCCCTTACAAAGACTAGATCTCTCAAACTACA 1391  
 QY 6045 TGAACCCCTCGTACCACTCGCTGTGTAAGCCTATTTAATACACCCCTCAGTGGCT 6104  
 DB 1392 TGAACCCCTCGTACCACTCGCTGTGTAAGCCTATTTAATACACCCCTCAGTGGCT 1451  
 QY 6105 CCATGAGTCTCGGCCCAAAACCTACTAATCTGTTGGATATGCTCCCTCGAATTCAR 6164  
 DB 1452 CCATGAGTCTCGGCCCAAAACCTACTAATCTGTTGGATATGCTCCCTCGAATTCAG 1511  
 QY 6165 GGCATATGTTTCAATCCCTGTACCTGAAACAATGGAACAACCTTCAGACAGAAATAAAC 6224  
 DB 1512 GGCATATGTTTCAATCCCTGTACCTGAAACAATGGAACAACCTTCAGACAGAAATAAAC 1571  
 QY 6225 CACTTCGGTTTGTAGTACGCTCTGTTTCCAAATGGAATTAACCACTACCTCAAACT 6284  
 DB 1572 CACTTCGGTTTGTAGTACGCTCTGTTTCCAAATGGAATTAACCACTACCTCAAACT 1631  
 QY 6285 CACTTCGTGTAATTTAGCAATACATACACCAACTCCCAATGATCAGGTGGGT 6344  
 DB 1632 CACTTCGTGTAATTTAGCAATACATACACCAACTCCCAATGATCAGGTGGGT 1691  
 QY 6345 AACTCTCCCAACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTC 6404  
 DB 1692 AACTCTCCCAACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTC 1751  
 QY 6405 AGCCTATCGTGTGTTGAATGGCTCTTCAAACTATGTCCTCTCAATCTTAGTGCC 6464  
 DB 1752 AGCCTATCGTGTGTTGAATGGCTCTTCAAACTATGTCCTCTCAATCTTAGTGCC 1811  
 QY 6465 CCCYATGRCATCTACACTGAAACAAGATTTATACAGTTATGTATATCTAAGCCCCGCA 6524  
 DB 1812 CCCYATGRCATCTACACTGAAACAAGATTTATACAGTTATGTATATCTAAGCCCCGCA 1871

QY 6525 CAAAGAGTACCATTCTTCTCTTTTATAGGAGCAGAGTGTAGCTGCAGTAGGTAC 6584  
 DB 1872 CAAAGAGTACCATTCTTCTCTTTTATAGGAGCAGAGTGTAGCTGCAGTAGGTAC 1931  
 QY 6585 TGGCATTGGCGGTATCAAACTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAA 6644  
 DB 1932 TGGCATTGGCGGTATCAAACTCTACTCAGTTCTTACTACAACTATCTCAAGAACTAAA 1991  
 QY 6645 TGGGACATGGAACGGGTGCGGACTCCCTGCTGCTGCAAGATCAACTTAATCCTCCT 6704  
 DB 1992 TGGGACATGGAACGGGTGCGGACTCCCTGCTGCTGCAAGATCAACTTAATCCTCCT 2051  
 QY 6705 AGCAGCAGTACTCTCTTC 6721  
 DB 2052 AGCAGCAGTACTCTCTTC 2068  
 RESULT 27  
 ID AAZ59468 standard; cDNA; 2946 BP.  
 XX  
 AC AAZ59468;  
 XX  
 DT 11-APR-2000 (first entry)  
 XX  
 DE Human secreted protein AJ172\_2 polynucleotide sequence.  
 XX  
 KW Human; secreted protein; disease diagnosis; pre-eclampsia; cancer;  
 KW placental pathology; metastasis inhibition; nutritional activity;  
 KW immune stimulator; haematopoiesis regulator; tissue growth;  
 KW tumour inhibitor; anti-inflammatory; clone AJ172\_2; ATCC\_98115;  
 KW gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9960020-A1.  
 XX  
 PD 25-NOV-1999.  
 XX  
 PF 17-MAY-1999; 99WO-US010915.  
 XX  
 PR 18-MAY-1998; 98US-00080478.  
 PR 20-OCT-1998; 98US-00175928.  
 XX  
 PA (GENY ) GENETICS INST INC.  
 XX  
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
 PI Merberg D, Mi S, Treacy M;  
 XX  
 DR WPI; 2000-116311/10.  
 DR P-PSDB; AAY67313.  
 XX  
 PT New polynucleotides encoding secreted cDNA libraries, used to develop  
 products for the diagnosis and treatment of neoplastic disease.  
 XX  
 PS Claim 14; Page 107-108; 149pp; English.  
 XX  
 CC This is the human secreted protein AJ172\_2 nucleotide sequence, obtained  
 from a human adult testes cDNA library. The invention relates to secreted  
 human and murine proteins. The polynucleotides and proteins are predicted  
 to have biological activities which would make them suitable for  
 treating, preventing or ameliorating medical conditions in humans and  
 animals. Detection of the levels of the proteins can be used for the  
 diagnosis of e.g. pre-eclampsia, placental pathology or cancer. Agents  
 which modulate the expression or function of the proteins may be used for  
 treating a neoplastic disease and inhibiting metastasis. Other suggested  
 activities include nutritional activity (e.g. in feeds), cytokine and  
 cell proliferation/differentiation activity, immune stimulating (e.g. as  
 vaccines) or suppressing activity, haematopoiesis regulating activity,  
 tissue growth activity, activin/inhibin activity,  
 chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,  
 receptor/ligand activity, anti-inflammatory activity, cadherin/tumour  
 invasion suppressor activity, and tumour inhibition activity. The

CC polynucleotide sequences are also stated to be useful for gene therapy  
 XX  
 SQ Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 U; 0 Other;  
 Query Match 9.0%; Score 680; DB 3; Length 2946;  
 Best Local Similarity 99.3%; Pred. No. 0;  
 Matches 1030; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5685 AGATTTCCTATGGAGATGCGGCTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCT 5744  
 |||||  
 Db 1032 AGATTTCCTATGGAGATGCGGCTCCCGGAAATATTGATGCCCCATCGTATAGGAGTCT 1091  
 |||||

QY 5745 TTSTAAGGAACCCCGACCTTCTAGTCCGACACCATATGCCGCAACTGCTATCACTC 5804  
 |||||  
 Db 1092 TTCTAAGGAACCCCGACCTTCTAGTCCGACACCATATGCCGCAACTGCTATCACTC 1151  
 |||||

QY 5805 TGGCCTCTTTGATGATGCAATATCTATTTATGACAGGAAATATGATTAATCTCTAG 5864  
 |||||  
 Db 1152 TGGCCTCTTTGATGATGCAATATCTATTTATGACAGGAAATATGATTAATCTCTAG 1211  
 |||||

QY 5865 TTGTCTCGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTACCCAAACTGCTATGTC 5924  
 |||||  
 Db 1212 TTGTCTCGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTACCCAAACTGCTATGTC 1271  
 |||||

QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAGAACTATCTCCCA 5984  
 |||||  
 Db 1272 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAGAACTATCTCCCA 1331  
 |||||

QY 5985 ACTCACGGGTGATCATGGCACTCTAGCCCTACAAAGGACTAGATCTCTCAAACTTACA 6044  
 |||||  
 Db 1332 ACTCACGGGTGATCATGGCACTCTAGCCCTACAAAGGACTAGATCTCTCAAACTTACA 1391  
 |||||

QY 6045 TGAACCTCTCGTACCCATCTGCTGCTGTAAGCTATTTTAAACCACTCTACTGGGCT 6104  
 |||||  
 Db 1392 TGAACCTCTCGTACCCATCTGCTGCTGTAAGCTATTTTAAACCACTCTACTGGGCT 1451  
 |||||

QY 6105 CCATGAGTCTCGGCCCAAAACCTTACTACTGTGATGCTGCTCCCTGCACTTTCAR 6164  
 |||||  
 Db 1452 CCATGAGTCTCGGCCCAAAACCTTACTACTGTGATGCTGCTCCCTGCACTTTCAR 1511  
 |||||

QY 6165 GCATATGTTTCAATCCCTGCTGCTGCAATGCAACCTTCTGAGCAAGAAATTAACAC 6224  
 |||||  
 Db 1512 GCATATGTTTCAATCCCTGCTGCTGCAATGCAACCTTCTGAGCAAGAAATTAACAC 1571  
 |||||

QY 6225 CACTTCGGTTTATGAGGACCTCTGTTTCCAAATSTGGAATTAACCATACCTCAAACT 6284  
 |||||  
 Db 1572 CACTTCGGTTTATGAGGACCTCTGTTTCCAAATSTGGAATTAACCATACCTCAAACT 1631  
 |||||

QY 6285 CACTGCTGTAATTTAGCAATACCTATACACCACTCCCAATGCAATCAGGTGGGT 6344  
 |||||  
 Db 1632 CACTGCTGTAATTTAGCAATACCTATACACCACTCCCAATGCAATCAGGTGGGT 1691  
 |||||

QY 6345 AACTCTCCCAACAATATGCTGCTTACCCTCAGGAATATTTTGTCTGTTGTTACCTC 6404  
 |||||  
 Db 1692 AACTCTCCCAACAATATGCTGCTTACCCTCAGGAATATTTTGTCTGTTGTTACCTC 1751  
 |||||

QY 6405 AGCCTATGTTGTTGAATGGCTCTTCAAGATCTATGCTGCTCTCTCATTTCTTAGTGCC 6464  
 |||||  
 Db 1752 AGCCTATGTTGTTGAATGGCTCTTCAAGATCTATGCTGCTCTCTCATTTCTTAGTGCC 1811  
 |||||

QY 6465 CCCATGRCATCTACACTGAACAGATTATATGATGATGATGATGATGATGATGATGATG 6524  
 |||||  
 Db 1812 CCCATGRCATCTACACTGAACAGATTATATGATGATGATGATGATGATGATGATGATG 1871  
 |||||

QY 6525 CAAAGAGTACCACTTCTCTTTTGTATAGGAGGAGGAGTCTAGGTGCACTAGGTAC 6584  
 |||||  
 Db 1872 CAAAGAGTACCACTTCTCTTTTGTATAGGAGGAGGAGTCTAGGTGCACTAGGTAC 1931  
 |||||

QY 6585 TGGCATTGGGGTATCACAACTCTACTCAGTTTCTACTACAACTATCTCAAGAACTAAA 6644  
 |||||  
 Db 1932 TGGCATTGGGGTATCACAACTCTACTCAGTTTCTACTACAACTATCTCAAGAACTAAA 1991  
 |||||

QY 6645 TGGGGACATGGAAACGGGTCCCGGACTCCCTGGTGCCTTGCAGATCAACTTAACCTCCCT 6704  
 |||||

Db 1992 TGGGGACATGGAAACGGGTCCCGGACTCCCTGGTGCACCTTGCAGATCAACTTAACCTCCCT 2051  
 |||||

QY 6705 AGCAGCAGTAGTCTCTTC 6721  
 |||||

Db 2052 AGCAGCAGTAGTCTCTTC 2068  
 |||||

RESULT 28  
 ADC38776  
 ID ADC38776 standard; cDNA; 2946 BP.  
 XX  
 AC ADC38776;  
 XX  
 DT 18-DEC-2003 (first entry)  
 XX  
 DE Human cDNA encoding a secreted protein #63.  
 XX  
 KW ss; gene; immune disorder; severe combined immunodeficiency; SCID;  
 KW autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;  
 KW rheumatoid arthritis; allergic reaction; asthma; myeloid cell deficiency;  
 KW lymphoid cell deficiency; osteoporosis; osteoarthritis;  
 KW peripheral nervous system disease; peripheral neuropathy;  
 KW Alzheimer's disease; Parkinson's disease; coagulation disorder;  
 KW inflammatory disease; systemic inflammatory response syndrome; SIRS;  
 KW ischaemia-reperfusion injury; Crohn's disease; anaphylaxis;  
 KW hypersensitivity; regeneration; neural cell proliferation; fertility;  
 KW tumour; chemokine; human; secreted protein.  
 XX  
 OS Homo sapiens.  
 XX  
 US2002193567-A1.  
 XX  
 PD 19-DEC-2002.  
 XX  
 PF 02-APR-2002; 2002US-00114893.  
 XX  
 PR 11-AUG-1995; 95US-00514014.  
 PR 05-APR-1996; 96US-00628364.  
 PR 19-APR-1996; 96US-00635311.  
 PR 07-JUN-1996; 96US-00659224.  
 PR 17-JUN-1996; 96US-00664596.  
 PR 09-JUL-1996; 96US-00677231.  
 PR 26-JUL-1996; 96US-00686878.  
 PR 23-AUG-1996; 96US-00701819.  
 PR 27-SEP-1996; 96US-00721488.  
 PR 27-SEP-1996; 96US-00721798.  
 PR 27-SEP-1996; 96US-00721923.  
 PR 27-SEP-1996; 96US-00721926.  
 PR 25-OCT-1996; 96US-00738367.  
 PR 30-OCT-1996; 96US-00739775.  
 PR 13-JAN-1997; 97US-00783395.  
 PR 10-APR-1997; 97US-00833823.  
 PR 02-JUN-1997; 97US-00867677.  
 PR 05-SEP-1997; 97US-00924838.  
 PR 06-OCT-1999; 99US-00413232.  
 XX  
 (GEMY ) GENETICS INST INC.  
 XX  
 PI Jacobs K, McCooy JM, Lavallie ER, Collins-Racie LA, Evans C;  
 PI Merberg D, Treacy M, Bowman MR, Spaulding V, Carlin-Duckett M;  
 PI Kelleher K;  
 XX  
 WIPI; 2003-657236/62.  
 DR P-PSDB; ADC38777.  
 XX  
 PT Proteins AZ3021 encoded by clone AZ3021 from human adult colon, and  
 PT BD12716 encoded by clone BD12716 from human fetal kidney cDNA library,  
 PT useful for treating e.g. multiple sclerosis and rheumatoid arthritis.  
 XX  
 PS Disclosure; SEQ ID NO 134; 412pp; English.  
 XX  
 CC The invention relates to a protein comprising fully defined AZ302 1

protein or BD127 1 6 protein. The polynucleotides are useful for expressing recombinant proteins for analysis and are also useful as chromosome markers or tags to identify chromosomes or to map related gene positions. The proteins are useful as amino acid supplement, carbon source, nitrogen source and carbohydrate source. The proteins are useful for treating various immune deficiencies and disorders (e.g. severe combined immunodeficiency (SCID)), autoimmune disorders (e.g. multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis), allergic reactions (e.g. asthma), myeloid or lymphoid cell deficiencies, osteoporosis or osteoarthritis, peripheral nervous system diseases (e.g. peripheral neuropathy, Alzheimer's disease, Parkinson's disease), coagulation disorders, inflammatory diseases (e.g. systemic inflammatory response syndrome (SIRS), ischaemia-reperfusion injury, Crohn's disease), anaphylaxis and hypersensitivity. Proteins are also useful for inducing tumour immunity, for inducing bone, cartilage, tendon, ligament and/or nerve growth or regeneration, for proliferating neural cells and for regenerating nerve and brain tissue, for inducing fertility and for inhibiting tumour growth. Proteins are also useful as chemokine for mammalian cells (e.g., monocytes, fibroblasts, neutrophils), and also useful as inhibitors of receptor/ligand interactions. The present sequence represents cDNA encoding a human secreted protein.

Query Match	9.0%;	Score 680;	DB 10;	Length 2946;
Best Local Similarity	99.3%;	Pred. No. 0;		
Matches 1030;	Conservative	0;	Mismatches	7; Indels 0; Gaps 0;
Qy	5685	AGAGTTTCTATGGAGATGACGGCTCCGGAAATATTGATGCCCATCGTATAGGAGTCT	5744	
Db	1032	AGAGTTTCTATGGAGAAATGACGGCTCCGGAAATATTGATGCCCATCGTATAGGAGTCT	1091	
Qy	5745	TTSTAAGGGAACCCCACTTCACTGCGCCACACCCATATGCCCGCAAATGCTCTATCACTC	5804	
Db	1092	TTCTAAGGGAACCCCACTTCACTGCGCCACACCCATATGCCCGCAAATGCTCTATCACTC	1151	
Qy	5805	TGCCACTCTTTGTCATGTCATCAATATTTGGACAGGAAAAATGATTAATCTCTAG	5864	
Db	1152	TGCCACTCTTTGTCATGTCATCAATATTTGGACAGGAAAAATGATTAATCTCTAG	1211	
Qy	5865	TTGTCTTGGAGGACTTGGAGTCACTGTCGTGGACTTACTTCAACCAACTGGTATGTC	5924	
Db	1212	TTGTCTTGGAGGACTTGGAGTCACTGTCGTGGACTTACTTCAACCAACTGGTATGTC	1271	
Qy	5925	TGATGGGGTGGAGTTCACAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCA	5984	
Db	1272	TGATGGGGTGGAGTTCACAGATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCA	1331	
Qy	5985	ACTCACSGGGTACATGGCACTCTAGCCCTACAAAGACTAGATCTCTCAAAACTTACA	6044	
Db	1332	ACTCACSGGGTACATGGCACTCTAGCCCTACAAAGACTAGATCTCTCAAAACTTACA	1391	
Qy	6045	TGAACCCCTCGTACCCATACCTGCGCTGGTAAGCTTAATTAACACCCCTCACTGGGCT	6104	
Db	1392	TGAACCCCTCGTACCCATACCTGCGCTGGTAAGCTTAATTAACACCCCTCACTGGGCT	1451	
Qy	6105	CCATGAGGTCTCGGCCCAAAACCCCTACTACTGTTGGATATGCTCCCGCTGAACTTCAR	6164	
Db	1452	CCATGAGGTCTCGGCCCAAAACCCCTACTACTGTTGGATATGCTCCCGCTGAACTTCAG	1511	
Qy	6165	GCCATATGTTTCAATCCCTGTACTGAAACAAATGGAAACAACTTCAGCACAGAAATAAACAC	6224	
Db	1512	GCCATATGTTTCAATCCCTGTACTGAAACAAATGGAAACAACTTCAGCACAGAAATAAACAC	1571	
Qy	6225	CAC TTCGTTTTAGTAGGA C C T T T G T T T C C A A T S T G G A A A T A A C C C A T A C C T C A A A C C T	6284	
Db	1572	CAC TTCGTTTTAGTAGGA C C T T T G T T T C C A A T S T G G A A A T A A C C C A T A C C T C A A A C C T	1631	
Qy	6285	CACCTGTGTAATAATTTAGCAATAC T A C A T A C A A C C A A C T C C C C A A T G C A T C A G G T G G G T	6344	
Db	1632	CACCTGTGTAATAATTTAGCAATAC T A C A T A C A A C C A A C T C C C C A A T G C A T C A G G T G G G T	1691	
Qy	6345	AAC TCCT C C C A C A C A A A T A G T C T G C C T A C C C T C A G G A A T A T T T T T G C T G T G G T A C C T C	6404	

Db	1692	AAC	TCCTCCACACAAATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGTACCTC	1751			
Qy	6405	AGC	TATCGTGTGTTTGAATGGCTCTTCAGAAATCTATGTGCTTCTCTCATTTCTTAGTGCC	6464			
Db	1752	AGC	TATCGTGTGTTTGAATGGCTCTTCAGAAATCTATGTGCTTCTCTCATTTCTTAGTGCC	1811			
Qy	6465	CCY	ATGRCATCTACACGTGAACAAGATTTATACAGTTATGTCAATCTTAAGCCCGCAA	6524			
Db	1812	CCCT	ATGACCATCTACACGTGAACAAGATTTATACAAATTTATGTCTATATCTAAGCCCGCAA	1871			
Qy	6525	CAA	AGATACCACTTCTTCTTTCTTATAGAGCAGGAGTGTAGGTGCATAGGTAC	6584			
Db	1872	CAA	AGAGTACCACTTCTTCTTTCTTATAGAGCAGGAGTGTAGGTGCATAGGTAC	1931			
Qy	6585	TGG	CATTGGCGGTATCACAACTCTACTCAGTTCTACTACAAAATCTCTCAAGAACTAAA	6644			
Db	1932	TGG	CATTGGCGGTATCACAACTCTACTCAGTTCTACTACAAAATCTCTCAAGAACTAAA	1991			
Qy	6645	TGG	GACATGGAACGGGTGCGGACTCCCTGGTCACTTGCAGAGTCAACTTAACTCCT	6704			
Db	1992	TGG	GACATGGAACGGGTGCGGACTCCCTGGTCACTTGCAGAGTCAACTTAACTCCT	2051			
Qy	6705	AGC	AGCAGTAGTCTCTTC	6721			
Db	2052	AGC	AGCAGTAGTCTCTTC	2068			
RESULT 29							
ADU81027							
XX	ADU81027	standard; DNA; 6332 BP.					
AC	ADU81027;						
XX	24-FEB-2005	(first entry)					
DT	Membrane-fusion protein containing viral vector, pCAGGS/Syncytin.						
DE	viral vector; enveloped virus; cytoplasmic; membrane-fusion;						
XX	pCAGGS/Syncytin; ds.						
KW	Cytomegalovirus.						
OS	Gallus gallus.						
OS	Oryctolagus cuniculus.						
OS	Simian virus 40.						
OS	Synthetic.						
XX							
FH	Key	Location/Qualifiers					
FT	enhancer	18..381					
FT		/*tag= a					
FT		/note= "Cytomegalovirus early enhancer"					
FT	promoter	382..662					
FT		/*tag= b					
FT		/note= "Chicken beta-actin promoter"					
FT	intron	663..1625					
FT		/*tag= c					
FT		/note= "Chicken beta-actin intron"					
FT	intron	1626..1670					
FT		/*tag= d					
FT		/note= "Rabbit beta-globin intron"					
FT	CDS	1725..3341					
FT		/*tag= e					
FT		/gene= "Syncytin gene"					
FT	misc_feature	3132..3338					
FT		/*tag= f					
FT		/note= "Cytoplasmic region"					
FT	polyA_signal	3359..3800					
FT		/*tag= g					
FT		/note= "Rabbit beta-globin poly A signal"					
FT	rep_origin	4002..4344					
FT		/*tag= h					
FT		/note= "SV40 replication origin"					
FT	CDS	5341..6201					





PR 15-SEP-1999; 99FR-00011793.  
 XX (INRM ) BIO MERIEUX.  
 PA (INRM ) INST NAT SANTE & RECH MEDICALE.  
 PI Mallet F, Cosset F, Blond J, Lavillette D, Bouton O, Ruggieri A;  
 XX  
 XX WPI; 2001-226676/23.  
 DR P-PSDB; AAB67652.  
 XX  
 PT Detecting expression of human endogenous retrovirus envelope protein in  
 PT cells of a tissue or culture, from its ability to induce syncytia.  
 XX  
 PS Disclosure; Page 44-45; 57pp; French.  
 XX  
 CC The present sequence encodes a human endogenous retrovirus envelope  
 CC protein. The specification describes a method for detecting expression of  
 CC an envelope protein from a human endogenous retrovirus (HERV), in cells,  
 CC of a tissue or culture. The method comprises detecting syncytia formation  
 CC due to the fusogenic properties of the envelope protein. Envelope  
 CC polypeptides and polynucleotides are used to produce therapeutic or  
 CC prophylactic compositions, particularly for treatment of cancer, to  
 CC correct defects in placental development (or other natural formation of  
 CC other types of syncytia), and to promote adhesion of cells in grafts or  
 CC cellular repair processes. Expression of sequences antisense to the  
 CC polynucleotide are used to prevent formation of syncytia  
 XX  
 SQ Sequence 2781 BP; 747 A; 770 C; 556 G; 708 T; 0 U; 0 Other;

Query Match 8.6%; Score 651; DB 5; Length 2781;  
 Best Local Similarity 99.3%; Pred. No. 0;  
 Matches 1001; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5685 AGAGTTTCTATGAGAAATCAGCGTCCCGGAAATATGATGCCCATCGTATAGGAGTCT 5744  
 DB 866 AGAGTTTCTATGAGAAATCAGCGTCCCGGAAATATGATGCCCATCGTATAGGAGTCT 925

QY 5745 TTSTAGGGAACCCACCTTCACTGCGCCACACCCCATATGCGCGGAACCTGCTATCCTC 5804  
 DB 926 TTCTAAGGGAACCCACCTTCACTGCGCCACACCCCATATGCGCGGAACCTGCTATCCTC 985

QY 5805 TGCCACTCTTTGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 5864  
 DB 986 TGCCACTCTTTGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1045

QY 5865 TTGTCTCTGAGGACTTTGGAGTCACTGCTGTTGGACTTACTTCAACCAAACTGGTATGTC 5924  
 DB 1046 TTGTCTCTGAGGACTTTGGAGTCACTGCTGTTGGACTTACTTCAACCAAACTGGTATGTC 1105

QY 5925 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAGAAGTAACTCTCCCA 5984  
 DB 1106 TGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAGAAGTAACTCTCCCA 1165

QY 5985 ACTCACCGGGTACATGGGACCTCTAGCCCTACAAAGACTAGATCTCTCAAAACTACA 6044  
 DB 1166 ACTCACCGGGTACATGGGACCTCTAGCCCTACAAAGACTAGATCTCTCAAAACTACA 1225

QY 6045 TGAAACCCCTCCGTACCCATCTCGCTTGGTAAAGCTTATTAATACCACTCACTGGGCT 6104  
 DB 1226 TGAAACCCCTCCGTACCCATCTCGCTTGGTAAAGCTTATTAATACCACTCACTGGGCT 1285

QY 6105 CCATGAGGTCTCGGCCCAAAACCTACTAACTGTTGGATATGCTCCCTGAACTTCAR 6164  
 DB 1286 CCATGAGGTCTCGGCCCAAAACCTACTAACTGTTGGATATGCTCCCTGAACTTCAG 1345

QY 6165 GCCATATGTTCAATCCCTGTACTGACAAATGGAACAACTTTCAGCAGCAAGAAATAACAC 6224  
 DB 1346 GCCATATGTTCAATCCCTGTACTGACAAATGGAACAACTTTCAGCAGCAAGAAATAACAC 1405

QY 6225 CACTTCCGTTTTAGTAGGACCTCTGTTTCCCAATSTGGAATAACCATACCTCAAACT 6284  
 DB 1406 CACTTCCGTTTTAGTAGGACCTCTGTTTCCCAATCTGGAAATAAACCATACCTCAAACT 1465

QY 6285 CACCTGTGTAATAATTTAGCAATACTACATACAAACCAACTCCCAATGCATCAGGTGGGT 6344  
 DB 1466 CACCTGTGTAATAATTTAGCAATACTACATACAAACCAACTCCCAATGCATCAGGTGGGT 1525

QY 6345 AACTCCTCCACACAAATAGTCTGCTTACCTCCAGGAATATTTTTTGTCTGTGGTACCTC 6404  
 DB 1526 AACTCCTCCACACAAATAGTCTGCTTACCTCCAGGAATATTTTTTGTCTGTGGTACCTC 1585

QY 6405 AGCCTATGTTGTTGAATGGCTCTTCAAGAAATCTATGTGCTTCTCTCATTTCTTAGTGCC 6464  
 DB 1586 AGCCTATGTTGTTGAATGGCTCTTCAAGAAATCTATGTGCTTCTCTCATTTCTTAGTGCC 1645

QY 6465 CCCYATGRCATCTACACTGAACAGATTTATACAGTTATGTCTATATCTAAGCCCGCAA 6524  
 DB 1646 CCCTATGACCATCTACACTGAACAGATTTATACAGTTATGTCTATATCTAAGCCCGCAA 1705

QY 6525 CAAAAGAGTACCCCATTTCTTCTTTTATAGGAGCAGGAGTGTAGGTGCACTAGGTAC 6584  
 DB 1706 CAAAAGAGTACCCCATTTCTTCTTTTATAGGAGCAGGAGTGTAGGTGCACTAGGTAC 1765

QY 6585 TGGCATTTGGCGGTATCAAACTCTCTACAGTTCTTACTACAAACTATCTCAAGAACTAAA 6644  
 DB 1766 TGGCATTTGGCGGTATCAAACTCTCTACAGTTCTTACTACAAACTATCTCAAGAACTAAA 1825

QY 6645 TGGGACATGGAACGGGTCCGCACTCCCTGCTGCTCACCCTTGCAGATCA 6692  
 DB 1826 TGGGACATGGAACGGGTCCGCACTCCCTGCTGCTCACCCTTGCAGATCA 1873

Search completed: August 2, 2006, 13:47:07  
 Job time : 4032 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2006, 15:08:55 ; Search time 27430 Seconds  
(without alignments)  
17675.878 Million cell updates/sec

Title: US-10-717-580-11  
Perfect score: 7489.6  
Sequence: 1 caacaatcgggataataacc.....tattaatacttgcaactgtr 7582

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenBank 152.\*

- 1: gb\_env.\*
- 2: gb\_pat.\*
- 3: gb\_ph.\*
- 4: gb\_pl.\*
- 5: gb\_pr.\*
- 6: gb\_ro.\*
- 7: gb\_sts.\*
- 8: gb\_sy.\*
- 9: gb\_un.\*
- 10: gb\_vi.\*
- 11: gb\_ov.\*
- 12: gb\_htg.\*
- 13: gb\_in.\*
- 14: gb\_om.\*
- 15: gb\_ba.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7489.6	100.0	7582	2	BD196245 Endogenet
2	7489.6	100.0	7582	2	BD267487 Endogenet
3	7489.6	100.0	7582	2	AX000966 Sequence
4	7489.6	100.0	7582	2	AX027480 Sequence
5	6448	86.1	10222	5	AY101582 Homo sapi
6	6448	86.1	10222	5	AY101585 Homo sapi
7	6448	86.1	56093	2	AX329572 Sequence
8	6448	86.1	56093	5	HSAC000064 Human BAC
9	6448	86.1	149194	5	AC007566 Homo sapi
10	6446.4	86.1	10222	5	AY101583 Homo sapi
11	6440	86.0	10222	5	AY101584 Homo sapi
12	6436	85.9	10499	2	BD221808 Nucleic s
13	6436	85.9	10499	2	AR699757 Sequence
14	6436	85.9	10499	2	AX007980 Sequence
15	6436.4	84.7	10229	5	AY101586 Pan trogl
16	6344.8	84.7	10229	5	AY101587 Pan trogl
17	6344.8	84.7	165590	5	AC161288 Pan trogl
18	6344.8	84.7	184675	5	AC145964 Pan trogl

C	19	6344.8	84.7	187270	5	AC161612 Pan trogl
	20	6269.6	83.7	10230	5	AY101588 Gorilla g
	21	6269.6	83.7	10230	5	AY101589 Gorilla g
	22	6124.6	81.8	8523	2	AR612522 Sequence
	23	6037.8	80.6	10122	5	AY101590 Pongo pyg
	24	6025.8	80.5	10124	5	AY101591 Pongo pyg
	25	5956.8	79.5	10246	5	AY101593 Hylobates
	26	5950.4	79.4	10248	5	AY101592 Hylobates
C	27	5824.2	77.8	158033	5	AC018926 Homo sapi
	28	5492	73.3	105989	5	AC008121 Homo sapi
	29	5492	73.3	110000	12	Continuation (3 of
	30	5380	71.8	169389	12	AC169739 Macaca mu
	31	5380	71.8	200397	12	AC144284 Macaca mu
C	32	4755.6	63.5	187321	12	AC092510 Papio anu
	33	4720.4	63.0	22272	5	AY925148 Macaca mu
	34	4302.2	57.4	114621	5	AC005187 Homo sapi
	35	4295.8	57.4	46575	12	AC080036 Homo sapi
	36	4162	55.6	111140	5	AL592310 Human DNA
	37	3944.6	52.7	183680	5	AC098859 Homo sapi
C	38	3779.8	50.5	73070	5	AL590143 Human DNA
	39	3731.6	49.8	173788	5	AC022555 Homo sapi
	40	3731.6	49.8	186540	5	AC090341 Homo sapi
	41	3731.6	49.8	187837	5	AC079065 Homo sapi
	42	3716.6	49.6	160883	12	AL158814 Homo sapi
C	43	3537.2	47.2	137947	5	HS45114 Human DNA
	44	3442.8	46.0	153444	12	AC040948 Homo sapi
	45	3442.8	46.0	205035	5	AL049870 Human chr

ALIGNMENTS

RESULT 1	BD196245	7582 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD196245	Endogenetic retroviral sequences, associated with autoimmune diseases and/or with pregnancy disorders.			
DEFINITION	BD196245	unclassified sequences.			
ACCESSION	BD196245.1	GI:33006015			
VERSION	JP 2002512530-A/11.				
KEYWORDS	unidentified				
SOURCE	unclassified				
ORGANISM	unclassified				
REFERENCE	1 (bases 1 to 7582)				
AUTHORS	Beseme,F., Blond,J.L., Bouton,O., Mandrand,B. and Mallet,F.				
TITLE	Endogenetic retroviral sequences, associated with autoimmune diseases and/or with pregnancy disorders				
JOURNAL	Patent: JP 2002512530-A 11 23-APR-2002;				
COMMENT	BIO MERIEUX				
	OS Unidentified				
	PN JP 2002512530-A/11				
	PD 23-APR-2002				
	PF 06-JUL-1998 JP 1999508244				
	PI 07-JUL-1997 FR 97/08815				
	PI FREDERIC BESEME, JEAN LUC BLOND, OLIVIER BOUTON, BERNARD MANDRAND,				
	PI FRANCOIS MALLET				
	PC C12N15/48,C07K14/15,C12Q1/68,C07K16/10,G01N33/569 CC				
	Strandedness: Single;				
	CC Topology: Linear;				
	CC Endogenetic retroviral sequences, associated with autoimmune diseases				
	CC and/or with pregnancy disorders				
	PH Key Location/Qualifiers				
	FT source				
		1..7582			
		Location/Qualifiers			
		1..7582			
		/organism="unidentified"			
		/mol_type="genomic DNA"			
		/db_xref="taxon:32644"			
FEATURES	source				
ORIGIN					

Query Match									
Best Local Similarity 100.0%; Score 7489.6; DB 2; Length 7582;									
Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	CAACAATCGGATATATAA	CCCAAGGCA	TTCCGAGCTGGCAAC	CAGCAGCCCCCTTTGGGTCC	60			
Db	1	CAACAATCGGATATATAA	CCCAAGGCA	TTCCGAGCTGGCAAC	CAGCAGCCCCCTTTGGGTCC	60			
Qy	61	CTTCCCTTTGTATGGAGCT	TTTTCA	TGCTATTTCACTCTATTAAT	TTGCAACTGCA	120			
Db	61	CTTCCCTTTGTATGGAGCT	TTTTCA	TGCTATTTCACTCTATTAAT	TTGCAACTGCA	120			
Qy	121	CTCTTCGTGCTCACTGTT	CTTACGGCTCGAGCTGAGCT	TTTCTC	CACCGTCCACCACTGC	180			
Db	121	CTCTTCGTGCTCACTGTT	CTTACGGCTCGAGCTGAGCT	TTTCTC	CACCGTCCACCACTGC	180			
Qy	181	TGTTTGGCACCAACCG	CANACCTG	CCGCTGACTCCCATCCCT	CTGGATCCTGCAAGGTGTC	240			
Db	181	TGTTTGGCACCAACCG	CANACCTG	CCGCTGACTCCCATCCCT	CTGGATCCTGCAAGGTGTC	240			
Qy	241	CGCTGTGCTCTGATCC	AGCARGCGCCCAT	TGCGGCTCCCAAT	TGGGCTAAAGGCTTGC	300			
Db	241	CGCTGTGCTCTGATCC	AGCARGCGCCCAT	TGCGGCTCCCAAT	TGGGCTAAAGGCTTGC	300			
Qy	301	CATTGTNCCTGCA	CGGCTAAGT	GCCTGGGTTTGTCTA	ATTGAGCTGAACACTANTCACT	360			
Db	301	CATTGTNCCTGCA	CGGCTAAGT	GCCTGGGTTTGTCTA	ATTGAGCTGAACACTANTCACT	360			
Qy	361	GGGTTCCATGGTCTCT	CTGTGACCCACGGCT	CTTAATAKAACTATA	TAACACTTACCACA	420			
Db	361	GGGTTCCATGGTCTCT	CTGTGACCCACGGCT	CTTAATAKAACTATA	TAACACTTACCACA	420			
Qy	421	TGGCCCAAGATTC	CAATTCCTTGAAAT	CCGTGAGGSCAACGAA	CTCAAGTCAAGAATAAC	480			
Db	421	TGGCCCAAGATTC	CAATTCCTTGAAAT	CCGTGAGGSCAACGAA	CTCAAGTCAAGAATAAC	480			
Qy	481	GARGCTTGCCACCAT	CTTTGGAAGCGGCT	GCTACBTTTGGAAGTGGT	TTCACCACTC	540			
Db	481	GARGCTTGCCACCAT	CTTTGGAAGCGGCT	GCTACBTTTGGAAGTGGT	TTCACCACTC	540			
Qy	541	TTGGGAGCTCTGTG	AGCAAGGACCCCGG	TRACATTTTGGCRAC	CCMSRACGACATCC	600			
Db	541	TTGGGAGCTCTGTG	AGCAAGGACCCCGG	TRACATTTTGGCRAC	CCMSRACGACATCC	600			
Qy	601	MAAGTATGGGAAAC	GGTTC	CCCGCAAGACAAA	ACGCCCTCAAGCTATTTCTGARAAT	660			
Db	601	MAAGTATGGGAAAC	GGTTC	CCCGCAAGACAAA	ACGCCCTCAAGCTATTTCTGARAAT	660			
Qy	661	TGGGAMCAATTTG	ACCTCAGACACT	TAAGAAAGAACGACT	TATATTTCTTGCAGTGCC	720			
Db	661	TGGGAMCAATTTG	ACCTCAGACACT	TAAGAAAGAACGACT	TATATTTCTTGCAGTGCC	720			
Qy	721	GCTGTGCACTCT	GTAGGGAAGTATAA	TTATAACCACTTTACAGCTAGAC	VTCTTTTG	780			
Db	721	GCTGTGCACTCT	GTAGGGAAGTATAA	TTATAACCACTTTACAGCTAGAC	VTCTTTTG	780			
Qy	781	TAGAAAAGGCAAT	GAGTGAAGTGC	CAATAAGTACAAA	CTTTCTTTTCAATTAAGAGACAA	840			
Db	781	TAGAAAAGGCAAT	GAGTGAAGTGC	CAATAAGTACAAA	CTTTCTTTTCAATTAAGAGACAA	840			
Qy	841	CTCACAATTTAT	GTAAAAAGTGA	TTTATGCCCCTACAGGAAGCTT	CAGAGTCTACCTCC	900			
Db	841	CTCACAATTTAT	GTAAAAAGTGA	TTTATGCCCCTACAGGAAGCTT	CAGAGTCTACCTCC	900			
Qy	901	CTATCCAGCAT	CCCCGACTCT	CTTCCC	CAATTAAGGACCCCTTCAACCCAAATGG	960			
Db	901	CTATCCAGCAT	CCCCGACTCT	CTTCCC	CAATTAAGGACCCCTTCAACCCAAATGG	960			
Qy	961	TCCAAAAGGAGAT	AGTAAACAGTGA	ACCAAGAGTGC	CAATATTTCCCAAT	1020			
Db	961	TCCAAAAGGAGAT	AGTAAACAGTGA	ACCAAGAGTGC	CAATATTTCCCAAT	1020			
Qy	1021	TATGACCCCTCC	CAAGCAGTGG	AGGAAGAATTCGG	CCCCCAGCAGATGTGCT	1080			

Db	1021	TATGACCCCTCC	CAAGCAGTGG	AGGAAGAATTCGG	CCCCCAGCAGATGTGCT	1080			
Qy	1081	TTTTYCTCC	CAGACTTAAAGCA	AATAAAAA	CAGACTTAGGTAAATTTCTCAGATAA	1140			
Db	1081	TTTTYCTCC	CAGACTTAAAGCA	AATAAAAA	CAGACTTAGGTAAATTTCTCAGATAA	1140			
Qy	1141	GATGGCTATAT	TGTGTTTAC	AGGGTTAGGAC	AAATTTCTTTGATCTCAGATGGAGAT	1200			
Db	1141	GATGGCTATAT	TGTGTTTAC	AGGGTTAGGAC	AAATTTCTTTGATCTCAGATGGAGAT	1200			
Qy	1201	ATATATGTCA	CTGCTTAAT	CAGACACTAA	CCCCCAATGAGAGAGTGC	1260			
Db	1201	ATATATGTCA	CTGCTTAAT	CAGACACTAA	CCCCCAATGAGAGAGTGC	1260			
Qy	1261	AGCCTGAGRT	TTGGCGATCT	CTGATCTCAGT	CAGTCAATGGATGATGACA	1320			
Db	1261	AGCCTGAGRT	TTGGCGATCT	CTGATCTCAGT	CAGTCAATGGATGATGACA	1320			
Qy	1321	GAAGGAAGANAAT	GATTTCC	CCACAGGCCAGC	AGTCCAGTCTASACCTCATTTG	1380			
Db	1321	GAAGGAAGANAAT	GATTTCC	CCACAGGCCAGC	AGTCCAGTCTASACCTCATTTG	1380			
Qy	1381	GGGACACAGAA	TCAAGTAA	CAATGGAGAT	TGGTCTGCAGACATTTGCTTA	1440			
Db	1381	GGGACACAGAA	TCAAGTAA	CAATGGAGAT	TGGTCTGCAGACATTTGCTTA	1440			
Qy	1441	TASAGGACTA	AGGAACTAS	AGAAATCTAY	GAAATTACTCAATGATGTCCACCATA	1500			
Db	1441	TASAGGACTA	AGGAACTAS	AGAAATCTAY	GAAATTACTCAATGATGTCCACCATA	1500			
Qy	1501	ACACAGGGAAG	GAAGAAATCT	TACTGCTG	GAGAGACTAAGGGAGGCAATTTG	1560			
Db	1501	ACACAGGGAAG	GAAGAAATCT	TACTGCTG	GAGAGACTAAGGGAGGCAATTTG	1560			
Qy	1561	GAAGCGTCC	CTCTGT	CACCTGACTCTT	CTGAAAGCCAACTAACTTTAAAGCGTAAGTT	1620			
Db	1561	GAAGCGTCC	CTCTGT	CACCTGACTCTT	CTGAAAGCCAACTAACTTTAAAGCGTAAGTT	1620			
Qy	1621	TATCACTCAG	TCAGTGC	AGACATTTAG	AAAAAACTTTCAAAAGTCTGCGCTAGGCGCGGAG	1680			
Db	1621	TATCACTCAG	TCAGTGC	AGACATTTAG	AAAAAACTTTCAAAAGTCTGCGCTAGGCGCGGAG	1680			
Qy	1681	CAAACTTTAG	AAACCTATT	CGAATTTG	GAACCTCGGTTTATAATAGATCAGGAG	1740			
Db	1681	CAAACTTTAG	AAACCTATT	CGAATTTG	GAACCTCGGTTTATAATAGATCAGGAG	1740			
Qy	1741	GAGCAGGCG	GAACAGCA	CAAAACCGGAT	TTAAAAAAGGCCACCGCTTTAGTCA	1800			
Db	1741	GAGCAGGCG	GAACAGCA	CAAAACCGGAT	TTAAAAAAGGCCACCGCTTTAGTCA	1800			
Qy	1801	CAGGCAAGT	GGACCTTTG	GAGGCTCTG	GAAAAAGGAAAAAGTGGGCAAAATTTGAATGCTTAA	1860			
Db	1801	CAGGCAAGT	GGACCTTTG	GAGGCTCTG	GAAAAAGGAAAAAGTGGGCAAAATTTGAATGCTTAA	1860			
Qy	1861	TAGGCTTGT	CTCCAGTGG	GTCTACAGG	CACTTTAAAAAGATTTGCTCAAGTAGAG	1920			
Db	1861	TAGGCTTGT	CTCCAGTGG	GTCTACAGG	CACTTTAAAAAGATTTGCTCAAGTAGAG	1920			
Qy	1921	TAAGCCGCC	CTTCTG	CTCAATGCC	CTTATTTCAAGGGAATCACTGGAAGGCCACCTG	1980			
Db	1921	TAAGCCGCC	CTTCTG	CTCAATGCC	CTTATTTCAAGGGAATCACTGGAAGGCCACCTG	1980			
Qy	1981	CAGGGAACAA	AGGTCTTTT	TGAGTCAG	AGGCCATAACAGATGATCCACGAGGAGCTG	2040			
Db	1981	CAGGGAACAA	AGGTCTTTT	TGAGTCAG	AGGCCATAACAGATGATCCACGAGGAGCTG	2040			
Qy	2041	AGGGTGC	CTGGGCAAG	CGCCATCC	ATGCCATCAAGAGCCCTGCGGTATGCTT	2100			
Db	2041	AGGGTGC	CTGGGCAAG	CGCCATCC	ATGCCATCAAGAGCCCTGCGGTATGCTT	2100			
Qy	2101	GACCAAT	TGAGGCG	CAGGAAGT	TGCTCTCTGACACTGGTGGGCTTCTT	2160			

Db 2101 GACCATTTAGGGCCAGGAAGGTTGTCTCTTGGACACTGGTGGGTCTTCTTAGTCTTACT 2160  
Qy 2161 CTTTCTGTCGGGACAACTGTCTCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGAGC 2220  
Db 2161 CTTTCTGTCGGGACAACTGTCTCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGAGC 2220  
Qy 2221 GGCAGTCACTAGATACCTTTTCCAGGACCACTAAGTTATGAACCTGGGAGCTTTATCTTTT 2280  
Db 2221 GGCAGTCACTAGATACCTTTTCCAGGACCACTAAGTTATGAACCTGGGAGCTTTATCTTTT 2280  
Qy 2281 TCACATGCTTTTCTAATATATGCTTTGAAAGCCCCACTACCTTGTGTAGGAGAGACATTTCTA 2340  
Db 2281 TCACATGCTTTTCTAATATATGCTTTGAAAGCCCCACTACCTTGTGTAGGAGAGACATTTCTA 2340  
Qy 2341 GCAAAAGCAGGGGCCATTATACACCTGAAATATAGGAGAGAAACACCCGTTGTGTGNNC 2400  
Db 2341 GCAAAAGCAGGGGCCATTATACACCTGAAATATAGGAGAGAAACACCCGTTGTGTGNNC 2400  
Qy 2401 CCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAGGACAATATGACGAGC 2460  
Db 2401 CCTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAGGACAATATGACGAGC 2460  
Qy 2461 CAAAGAAATGCCGTCCTGTTCAAAGTTAAACTAAAGGATTCACCTTCTTCCCTACCAA 2520  
Db 2461 CAAAGAAATGCCGTCCTGTTCAAAGTTAAACTAAAGGATTCACCTTCTTCCCTACCAA 2520  
Qy 2521 GGCAGTACCCCTCAGACCCCAAGGCCCAACAGAGATTCCAAAGATTTGTAAGGACTTAA 2580  
Db 2521 GGCAGTACCCCTCAGACCCCAAGGCCCAACAGAGATTCCAAAGATTTGTAAGGACTTAA 2580  
Qy 2581 AAGCCCAAGGCTTAGTAAACCAATGATTAACCTCCCTGCAAGTATTCGTTAGTGGATGAG 2640  
Db 2581 AAGCCCAAGGCTTAGTAAACCAATGATTAACCTCCCTGCAAGTATTCGTTAGTGGATGAG 2640  
Qy 2641 GAGGCA CAGAAACCCAGTGGACAGTGGAGGTTAGTCAAGATCTCAGGATTAACAATGG 2700  
Db 2641 GAGGCA CAGAAACCCAGTGGACAGTGGAGGTTAGTCAAGATCTCAGGATTAACAATGG 2700  
Qy 2701 AGGCGTGTGCTTTTATACCAAGCTGTACTAGCCCTTATATCTGTMVTTCCCAATAC 2760  
Db 2701 AGGCGTGTGCTTTTATACCAAGCTGTACTAGCCCTTATATCTGTMVTTCCCAATAC 2760  
Qy 2761 CAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCTTCTTCTGATCCCTG 2820  
Db 2761 CAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAGGATGCTTCTTCTGATCCCTG 2820  
Qy 2821 TACATCCTGACTCTCAATCTTGTGTCCTTTGAAGATCTCAAACCCCAACATCTCAAC 2880  
Db 2821 TACATCCTGACTCTCAATCTTGTGTCCTTTGAAGATCTCAAACCCCAACATCTCAAC 2880  
Qy 2881 TCACCTGGACTRTTTTACCCCAAGGTTTCAAGGATAGYCCCCATCTATTTTGGCCAGGCAT 2940  
Db 2881 TCACCTGGACTRTTTTACCCCAAGGTTTCAAGGATAGYCCCCATCTATTTTGGCCAGGCAT 2940  
Qy 2941 TAGCCCAAGACTTGAGYCACTYMTFCATCTGACACTCTGTCTTCTGRTAGTGGATG 3000  
Db 2941 TAGCCCAAGACTTGAGYCACTYMTFCATCTGACACTCTGTCTTCTGRTAGTGGATG 3000  
Qy 3001 ATTTTACTTTTTRGCGCYRTTTCAGAAACCTTTGTCATCAAGCCACCCAGGCTCTTWA 3060  
Db 3001 ATTTTACTTTTTRGCGCYRTTTCAGAAACCTTTGTCATCAAGCCACCCAGGCTCTTWA 3060  
Qy 3061 ATTTTCTCGCYACCTGTGGCTACAWGTTTTCAAAGSABARGCTCACTCTGCTCACAGC 3120  
Db 3061 ATTTTCTCGCYACCTGTGGCTACAWGTTTTCAAAGSABARGCTCACTCTGCTCACAGC 3120  
Qy 3121 AGGTTAAATCTTAGGCTTAARATTTTCCAAAGGCAACCAAGGCCCTCAGTGAGGAAYRYA 3180  
Db 3121 AGGTTAAATCTTAGGCTTAARATTTTCCAAAGGCAACCAAGGCCCTCAGTGAGGAAYRYA 3180  
Qy 3181 TCCAGCCTTACTGGCTTATCCTCATCYCAAAACCTTAAGCAACTAAGGRRRTTCTTGG 3240  
Db 3181 TCCAGCCTTACTGGCTTATCCTCATCYCAAAACCTTAAGCAACTAAGGRRRTTCTTGG 3240

Qy 3241 GCRTAAYAGGYTTCTCCCAAWATGATTTCCAGGTTWGGCRAAATAGCCAGGYCATTA 3300  
Db 3241 GCRTAAYAGGYTTCTCCCAAWATGATTTCCAGGTTWGGCRAAATAGCCAGGYCATTA 3300  
Qy 3301 WATACASTAATTAAGGAAACTCAGAAAGCCAAATACCCATTTARTTAAGATGGAAYMCTGAA 3360  
Db 3301 WATACASTAATTAAGGAAACTCAGAAAGCCAAATACCCATTTARTTAAGATGGAAYMCTGAA 3360  
Qy 3361 GYMRAAGTGGCTTTCCAGGCCCTTAAAGAGGCTTTAAACCCCAAGYCCCAGTGTTAAGYT 3420  
Db 3361 GYMRAAGTGGCTTTCCAGGCCCTTAAAGAGGCTTTAAACCCCAAGYCCCAGTGTTAAGYT 3420  
Qy 3421 TGCCAA CRGGGCAAGACTTTTSTTYATATRTCAAGAAAAAACAAGAAAYAGCTCTRGGG 3480  
Db 3421 TGCCAA CRGGGCAAGACTTTTSTTYATATRTCAAGAAAAAACAAGAAAYAGCTCTRGGG 3480  
Qy 3481 GTCTCTTACACAGTCCCRAGGGAYGAGCTTGCACACCTVGTGCRVACCTGASTAGGAAAYT 3540  
Db 3481 GTCTCTTACACAGTCCCRAGGGAYGAGCTTGCACACCTVGTGCRVACCTGASTAGGAAAYT 3540  
Qy 3541 GATGTAGTGGCAAGGTTGRCVYTCATTTGTAYGGTGTAGTGGTGGCAGTAGCAGTYKTA 3600  
Db 3541 GATGTAGTGGCAAGGTTGRCVYTCATTTGTAYGGTGTAGTGGTGGCAGTAGCAGTYKTA 3600  
Qy 3601 GTATCTGAAGCAGTTAAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
Db 3601 GTATCTGAAGCAGTTAAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
Qy 3661 GTGAAYRGCATACTCACTCTAAGGAGACTTGTGGCTGTCAAGAACVGTTTACTTAA 3720  
Db 3661 GTGAAYRGCATACTCACTCTAAGGAGACTTGTGGCTGTCAAGAACVGTTTACTTAA 3720  
Qy 3721 TTTCCAGGCTCTATTACTTGAARGGCCAGTGTGCTGCACTGTGCACTTGTGCAACTCTTAAC 3780  
Db 3721 TTTCCAGGCTCTATTACTTGAARGGCCAGTGTGCTGCACTGTGCACTTGTGCAACTCTTAAC 3780  
Qy 3781 CCAGYCNCAATTTCTTCCAGCAATGAAGAAAAAGATARAAYATAACTGTCAACAARTAAAT 3840  
Db 3781 CCAGYCNCAATTTCTTCCAGCAATGAAGAAAAAGATARAAYATAACTGTCAACAARTAAAT 3840  
Qy 3841 TCTCAAAACCTATGCCACTCGAGGGACCTTGTAGAGTTCCVTTGACTGATCCYAGACCTT 3900  
Db 3841 TCTCAAAACCTATGCCACTCGAGGGACCTTGTAGAGTTCCVTTGACTGATCCYAGACCTT 3900  
Qy 3901 CAACTTGATATCTGATGGAAGTTCTTTGTAGAAAAGGACTTTCGAAAAGYGGGGTATGC 3960  
Db 3901 CAACTTGATATCTGATGGAAGTTCTTTGTAGAAAAGGACTTTCGAAAAGYGGGGTATGC 3960  
Qy 3961 AGTGTGCTAGTGAATAATGGAATAYTTGAAAGTAATCCCTCACTCCAGGAACCTAGTCTYA 4020  
Db 3961 AGTGTGCTAGTGAATAATGGAATAYTTGAAAGTAATCCCTCACTCCAGGAACCTAGTCTYA 4020  
Qy 4021 GCTRCAGAACTAATAGCCYTCAYTKGGCACTAGAAATTAGGAGAAAGRAAAAAGGGYAAA 4080  
Db 4021 GCTRCAGAACTAATAGCCYTCAYTKGGCACTAGAAATTAGGAGAAAGRAAAAAGGGYAAA 4080  
Qy 4081 TATATATACAGACTCTRTATATGCTYACCTAGTCTCCATGCCATGMRGCAATAGSAR 4140  
Db 4081 TATATATACAGACTCTRTATATGCTYACCTAGTCTCCATGCCATGMRGCAATAGSAR 4140  
Qy 4141 AGAAGGGGAATTCCTAACTTCYAGRGAAACCTATCAMACATCAGGAAGCCATTAGGAR 4200  
Db 4141 AGAAGGGGAATTCCTAACTTCYAGRGAAACCTATCAMACATCAGGAAGCCATTAGGAR 4200  
Qy 4201 ATTATATYTGCGWGTACAGAAACCTTARAGAGTGTGMAGTCTTTACACTGCGGGGTATCA 4260  
Db 4201 ATTATATYTGCGWGTACAGAAACCTTARAGAGTGTGMAGTCTTTACACTGCGGGGTATCA 4260  
Qy 4261 NAAAGGAAGRAAAGGGAAATASAGRNAAYTGCCCAAGCAKATATTTAGGCMWAAGAGC 4320  
Db 4261 NAAAGGAAGRAAAGGGAAATASAGRNAAYTGCCCAAGCAKATATTTAGGCMWAAGAGC 4320

QY 4321 TCGAAGGAGGAGCCCTCCATTAGAAATGCTTATTAACCTTCCTTAGTATAGGGTAATCC 4380  
 DB 4321 TCGAAGGAGGAGCCCTCCATTAGAAATGCTTATTAACCTTCCTTAGTATAGGGTAATCC 4380  
 QY 4381 CTTTCGGGAAACCAAGCCCTCAGTACTCAGCAGGAGAAACAGAAATGGGAAACCTCAGAGG 4440  
 DB 4381 CTTTCGGGAAACCAAGCCCTCAGTACTCAGCAGGAGAAACAGAAATGGGAAACCTCAGAGG 4440  
 QY 4441 CAGTTTTTCTCCCTCGGGACGGTGTAGCCACTGAAAGAGGAAATACCTTTGCTGCAAC 4500  
 DB 4441 CAGTTTTTCTCCCTCGGGACGGTGTAGCCACTGAAAGAGGAAATACCTTTGCTGCAAC 4500  
 QY 4501 TATCAATGGAATTTACTTTAAACCTTCATCAACCTTTTCACTTTAGGCATCGATAGCAC 4560  
 DB 4501 TATCAATGGAATTTACTTTAAACCTTCATCAACCTTTTCACTTTAGGCATCGATAGCAC 4560  
 QY 4561 CCATCARATGGCCAAATCATTTATTTACTGGACAGGCGCTTTTCAAACTATCAAGCARAT 4620  
 DB 4561 CCATCARATGGCCAAATCATTTATTTACTGGACAGGCGCTTTTCAAACTATCAAGCARAT 4620  
 QY 4621 AKTCAGGCGCTGTGAATGTGCAARAAATTAATCCCTGCTCTYATGCGCAAGCTCCTTC 4680  
 DB 4621 AKTCAGGCGCTGTGAATGTGCAARAAATTAATCCCTGCTCTYATGCGCAAGCTCCTTC 4680  
 QY 4681 AGGARAACAAAPACAGGCCATTACCCTGRARAARACTGGCAACTGATTTTACCCACAAG 4740  
 DB 4681 AGGARAACAAAPACAGGCCATTACCCTGRARAARACTGGCAACTGATTTTACCCACAAG 4740  
 QY 4741 CCCAAACCTCAGGGATTTTCACTATCTACTAGTCTGGGTARATACCTTTTCAAGGGTGGCA 4800  
 DB 4741 CCCAAACCTCAGGGATTTTCACTATCTACTAGTCTGGGTARATACCTTTTCAAGGGTGGCA 4800  
 QY 4801 RAGGCGCTCCCTGTAGGACAGAAAGCCCAAGAGGTAAATAAGGCACTAGTTTCATGAA 4860  
 DB 4801 RAGGCGCTCCCTGTAGGACAGAAAGCCCAAGAGGTAAATAAGGCACTAGTTTCATGAA 4860  
 QY 4861 ATAAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTCCAG 4920  
 DB 4861 ATAAATCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTCCAG 4920  
 QY 4921 GCCACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTCGCGCC 4980  
 DB 4921 GCCACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTCGCGCC 4980  
 QY 4981 TGAAGCCACAGTCTCTCAGGAGAGTTCGAGAAATGAATGAAYACTCAAGGACATCTA 5040  
 DB 4981 TGAAGCCACAGTCTCTCAGGAGAGTTCGAGAAATGAATGAAYACTCAAGGACATCTA 5040  
 QY 5041 AAAAGCAACCCAGGAAACCCACCTCACATGGCCTGCTCTGCTATAGCCTTAAA 5100  
 DB 5041 AAAAGCAACCCAGGAAACCCACCTCACATGGCCTGCTCTGCTATAGCCTTAAA 5100  
 QY 5101 AGAATCTGCAATTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGG 5160  
 DB 5101 AGAATCTGCAATTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGG 5160  
 QY 5161 CCCTTCATACCAATGACCTGTGCTTGACCCCAAGACGCCAATTTAGTTGACGACATCA 5220  
 DB 5161 CCCTTCATACCAATGACCTGTGCTTGACCCCAAGACGCCAATTTAGTTGACGACATCA 5220  
 QY 5221 CCTCCTTAGCCAAATATCAAGTTCTTTAAACATTTACAGGAACCTTATCCCTGAGAAG 5280  
 DB 5221 CCTCCTTAGCCAAATATCAAGTTCTTTAAACATTTACAGGAACCTTATCCCTGAGAAG 5280  
 QY 5281 AGGGAAGAACTATTTCACCCWGTGACATGATTTAGTCAAGTCCCTTCTCTTAAT 5340  
 DB 5281 AGGGAAGAACTATTTCACCCWGTGACATGATTTAGTCAAGTCCCTTCTCTTAAT 5340  
 QY 5341 CCCCATCCCTAGATACATCTCGGAAGGACCCCTACCCAGTCAATTTTATYACCCCACTG 5400  
 DB 5341 CCCCATCCCTAGATACATCTCGGAAGGACCCCTACCCAGTCAATTTTATYACCCCACTG 5400  
 QY 5401 CGGTTAAAGTGGCTGGAGTCTTTGGATACATCACACTTGAGTCAATCCTGGATAC 5460

DB 5401 CGGTTAAAGTGGCTGGAGTCTTTGGATACATCACACTTGAGTCAAAATCCTGGATAC 5460  
 QY 5461 TGCCAAAGGAACCTGAAATATCCAGAGACAACGCTAGCTATTTCTGTGAAACCTCTAGAGG 5520  
 DB 5461 TGCCAAAGGAACCTGAAATATCCAGAGACAACGCTAGCTATTTCTGTGAAACCTCTAGAGG 5520  
 QY 5521 ATTTGCGCTCTCTTCAAAACAACAGAGAGAAAGTAATAAATCATATAATATCCCCC 5580  
 DB 5521 ATTTGCGCTCTCTTCAAAACAACAGAGAGAAAGTAATAAATCATATAATATCCCCC 5580  
 QY 5581 ATGSGCCTCCCTTATCATATTTTCTCTKTASTGTTTTCCTTACCTCTCTCACT 5640  
 DB 5581 ATGSGCCTCCCTTATCATATTTTCTCTKTASTGTTTTCCTTACCTCTCTCACT 5640  
 QY 5641 GCACCCCCTCATGCGCTGTATGACAGTAGTCCCTCTYACMAGAGTTTCTATGGAGA 5700  
 DB 5641 GCACCCCCTCATGCGCTGTATGACAGTAGTCCCTCTYACMAGAGTTTCTATGGAGA 5700  
 QY 5701 ATGCAAGCGTCCGGAATATTTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCC 5760  
 DB 5701 ATGCAAGCGTCCGGAATATTTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCC 5760  
 QY 5761 ACCTTCACTGCGCCACACCCATATGCGCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820  
 DB 5761 ACCTTCACTGCGCCACACCCATATGCGCGCAACTGCTATCACTCTGCCACTCTTTGCATG 5820  
 QY 5821 CATGCAATATCTCATTTATTTGACAGGAAAAATGATTAATCTCTAGTTGCTTGGAGGACTT 5880  
 DB 5821 CATGCAATATCTCATTTATTTGACAGGAAAAATGATTAATCTCTAGTTGCTTGGAGGACTT 5880  
 QY 5881 GGAGTCACTGCTGTGTGAGCTTACTTCAACCCAACTGGTATGCTGTGATGGGGTGGAGTT 5940  
 DB 5881 GGAGTCACTGCTGTGTGAGCTTACTTCAACCCAACTGGTATGCTGTGATGGGGTGGAGTT 5940  
 QY 5941 CAAGATCAGGCAAGAGAAAAACATGTAAGAAAGTAATCTCCAACTCACCCSGGGTACAT 6000  
 DB 5941 CAAGATCAGGCAAGAGAAAAACATGTAAGAAAGTAATCTCCAACTCACCCSGGGTACAT 6000  
 QY 6001 GGCACTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCGGTACC 6060  
 DB 6001 GGCACTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCGGTACC 6060  
 QY 6061 CATCTCGCTGTGTAAGCTATTTAATACCCCTCTACCTGGGCTCCATGAGTCTCGGCC 6120  
 DB 6061 CATCTCGCTGTGTAAGCTATTTAATACCCCTCTACCTGGGCTCCATGAGTCTCGGCC 6120  
 QY 6121 CAATAACCTTACTAACTGTGGATATGCTCCCTGAACTTCARGCCATATGTTTCAATC 6180  
 DB 6121 CAATAACCTTACTAACTGTGGATATGCTCCCTGAACTTCARGCCATATGTTTCAATC 6180  
 QY 6181 CCTGTAACCTGAAACAACTTACAGCAGAGAAATAAACAACCACTTCCGTTTTAGTA 6240  
 DB 6181 CCTGTAACCTGAAACAACTTACAGCAGAGAAATAAACAACCACTTCCGTTTTAGTA 6240  
 QY 6241 GGAACCTTGTTCCTTCCAAATSTGGAATAACCCATACCTCAAACTCACCTGTGTAAATTT 6300  
 DB 6241 GGAACCTTGTTCCTTCCAAATSTGGAATAACCCATACCTCAAACTCACCTGTGTAAATTT 6300  
 QY 6301 AGCAATCTACATACACAACCACTCCCAATGTCATCAGTGGTAACTCTCCCAACACAA 6360  
 DB 6301 AGCAATCTACATACACAACCACTCCCAATGTCATCAGTGGTAACTCTCCCAACACAA 6360  
 QY 6361 ATAGTCTGCTTACCCCTCAGGAATAATTTTGTCTGTGTACCTCAGCCCTATCGTTGTTG 6420  
 DB 6361 ATAGTCTGCTTACCCCTCAGGAATAATTTTGTCTGTGTACCTCAGCCCTATCGTTGTTG 6420  
 QY 6421 AATGCTCTTCAGAACTATGTGCTTCTCTCAATTTTAGTCCCTCATGRCATCTAC 6480  
 DB 6421 AATGCTCTTCAGAACTATGTGCTTCTCTCAATTTTAGTCCCTCATGRCATCTAC 6480  
 QY 6481 ACTGAAACAAGATTTATACAGTTATGTCTATCTAAAGCCCGCAACAAAGAGTACCAAT 6540



6481 ACTGAACAAGATTTATACAGTTATGTATCATATCTAAGCCCCCGCAACAAAAAGAGTACCCATT 6540  
6541 CTTTCCTTTTGTATAGAGCAGGAGTGTAGTGCACTAGTACTGCGATTGGCGGTATC 6600  
6541 CTTTCCTTTTGTATAGAGCAGGAGTGTAGTGCACTAGTACTGCGATTGGCGGTATC 6600  
6601 ACAACCTCTACTCAGTTTCTACTCAAACTATCTCAAGAACTAAATGGGACATGGAACGG 6660  
6601 ACAACCTCTACTCAGTTTCTACTCAAACTATCTCAAGAACTAAATGGGACATGGAACGG 6660  
6661 GTGCCCGACCTCCCTGGTCACTTGCAGAGTCAACTTAACTCCCTAGCAGCAGTAGTCCCT 6720  
6661 GTGCCCGACCTCCCTGGTCACTTGCAGAGTCAACTTAACTCCCTAGCAGCAGTAGTCCCT 6720  
6721 CBAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAGAGAGGGGAACTGTATTATTTTA 6780  
6721 CBAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAGAGAGGGGAACTGTATTATTTTA 6780  
6781 GGGGAAGATGCTGTATTATTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAGAA 6840  
6781 GGGGAAGATGCTGTATTATTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAGAA 6840  
6841 ATTCSAGATCGAATACAACTGAGCAGAGAGCTTCAAAACACTGAGACCTCGGGCCTC 6900  
6841 ATTCSAGATCGAATACAACTGAGCAGAGAGCTTCAAAACACTGAGACCTCGGGCCTC 6900  
6901 CTAGCCATGAGATGAGCTGCTGATTTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATG 6960  
6901 CTAGCCATGAGATGAGCTGCTGATTTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATG 6960  
6961 CTACTCCTCTTGGACCTGATCTTTTACCTCTCTTTTAACTTTGTTCTCTTCCAGATC 7020  
6961 CTACTCCTCTTGGACCTGATCTTTTACCTCTCTTTTAACTTTGTTCTCTTCCAGATC 7020  
7021 GAAGCTGTAAACTCAAAATGGAGCCCAAGATGAGTCAAGACTCAAGACTTAAGACTTACCAG 7080  
7021 GAAGCTGTAAACTCAAAATGGAGCCCAAGATGAGTCAAGACTCAAGACTTAAGACTTACCAG 7080  
7081 CCCTGAGACCGGCTGTAGCCCAAGATGAGTCAAGACTCAAGACTTAAGACTTACCAG 7140  
7081 CCCTGAGACCGGCTGTAGCCCAAGATGAGTCAAGACTCAAGACTTAAGACTTACCAG 7140  
7141 GAGGAATCTCAGCTGCAACACTCTACTACGCCCCCAATTCAGCAGGAGCAGTTAGAC 7200  
7141 GAGGAATCTCAGCTGCAACACTCTACTACGCCCCCAATTCAGCAGGAGCAGTTAGAC 7200  
7201 GGTSGTCGGCAACCTCCCAACAGCACTTAGGTTTCTGTTGAGATGGGACACTGAGA 7260  
7201 GGTSGTCGGCAACCTCCCAACAGCACTTAGGTTTCTGTTGAGATGGGACACTGAGA 7260  
7261 GACAGACTAGCTGGATTTCTTAGGCTGATTAAGAAATCCYTAAGCTTAGTGGGAAGGTG 7320  
7261 GACAGACTAGCTGGATTTCTTAGGCTGATTAAGAAATCCYTAAGCTTAGTGGGAAGGTG 7320  
7321 ACCACATCCACCTTTAAACAGCGGGCTTGCACTTAGTTCACACCTGACCAATCAGAGAG 7380  
7321 ACCACATCCACCTTTAAACAGCGGGCTTGCACTTAGTTCACACCTGACCAATCAGAGAG 7380  
7381 CTCACCTAAATGCTAATTAGGCAAGACAGGAGTAAAGAAATAGCCATCATYTTATTC 7440  
7381 CTCACCTAAATGCTAATTAGGCAAGACAGGAGTAAAGAAATAGCCATCATYTTATTC 7440  
7441 MTGAGAGCAGCAGGAGGAGCAATGATCGGGATATAAACCCCAAGTTCGAGCCCGGCA 7500  
7441 MTGAGAGCAGCAGGAGGAGCAATGATCGGGATATAAACCCCAAGTTCGAGCCCGGCA 7500  
7501 CGGCAACCCCTTTGGGTCCCTCTTGTATGGGAGCTCTGTTTTCATGCTATTTCAC 7560  
7501 CGGCAACCCCTTTGGGTCCCTCTTGTATGGGAGCTCTGTTTTCATGCTATTTCAC 7560  
7561 TCTATTAAATCTTGCARCTGCR 7582  
7561 TCTATTAAATCTTGCARCTGCR 7582

RESULT 2  
BD267487  
LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## PD

## PF

## PI

## G12N15/09

## G01N33/564

## PC

## CC

## FH

## FT

## FEATURES

## source

## ORIGIN

## Query Match

## Best Local Similarity

## Matches 7582;

## QY

## Db

## QY

## Db

## QY

## Db

## QY

## Db

## QY

## Db

## QY

## Db

## QY

## Db

BD267487 7582 bp DNA linear PAT 17-JUL-2003  
Endogenous nucleic fragment associated with an autoimmune disease,  
marking method and reagent.

BD267487  
BD267487.1 GI:33077255  
JP 2002534980-A/28.  
Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

Baccala, G.P., Mallet, P. and Voisset, C.  
Endogenous nucleic fragment associated with an autoimmune disease,  
marking method and reagent

Patent: JP 2002534980-A 28 22-OCT-2002;  
BIO MERIEUX  
OS Homo sapiens (human)

PN JP 2002534980-A/28  
PD 22-OCT-2002  
PF 21-JAN-2000 JP 2000594929

PI GLAUCIA PARANHOS BACCALA, FRANCOIS MALLET, CECILE VOISSET PC  
G12N15/09, G07K14/15, C12Q1/02, C12Q1/68, G01N33/53, G01N33/564

PC G01N33/566, G01N33/569, G01N33/58//C12P21/02, C12N15/00 CC  
Endogenous nucleic fragment associated with an autoimmune CC  
disease,

CC marking method and reagent  
FH Key Location/Qualifiers  
FT source 1..7582  
FT /organism='Homo sapiens (human)'

source 1..7582  
/organism='Homo sapiens'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:9606'

Query Match 100.0%; Score 7489.6; DB 2; Length 7582;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAACATCGGATATAAACCAGGCAATTCAGCTGGCAACAGAGCCCCCTTTGGGTCC 60  
Db 1 CAACATCGGATATAAACCAGGCAATTCAGCTGGCAACAGAGCCCCCTTTGGGTCC 60

QY 61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTTCACTCTATTAATCTTGCAACTGCA 120  
Db 61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTTCACTCTATTAATCTTGCAACTGCA 120

QY 121 CTCCTTCGCTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTACCGCTCCACCTGTC 180  
Db 121 CTCCTTCGCTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTACCGCTCCACCTGTC 180

QY 181 TGTGTTGCCACCAACCGCANACCTGCGCTGACTCCCATCCCTCTGGATCTTCGAGGGTGC 240  
Db 181 TGTGTTGCCACCAACCGCANACCTGCGCTGACTCCCATCCCTCTGGATCTTCGAGGGTGC 240

QY 241 CGCTGTGCTCCTGATCCAGAGCGCCCATTTGCCGCTCCCAATTTGGGTAAAGGCTTCG 300  
Db 241 CGCTGTGCTCCTGATCCAGAGCGCCCATTTGCCGCTCCCAATTTGGGTAAAGGCTTCG 300

QY 301 CATTTGNCCTGCAACCGCTGAAGTGCCTGGGTTTGTCTTAATTCAGCTGAACACTANTCACT 360  
Db 301 CATTTGNCCTGCAACCGCTGAAGTGCCTGGGTTTGTCTTAATTCAGCTGAACACTANTCACT 360

QY 361 GGGTTCATGTTCTCTCTGTGACCCACCGCTTCTTAATAKAACTATAACACTTACCACA 420  
Db 361 GGGTTCATGTTCTCTCTGTGACCCACCGCTTCTTAATAKAACTATAACACTTACCACA 420

Db 361 GGGTCCATGGTCTCTCTGTGACCCACGGCTTCTAATAKAACTATAACACTTACCACA 420  
Qy 421 TGGCCCAAGATTCCATTCCTTGGGAATCCGTGAGGSAACGAACCTCCAGGTGAGAAATAC 480  
Db 421 TGGCCCAAGATTCCATTCCTTGGGAATCCGTGAGGSAACGAACCTCCAGGTGAGAAATAC 480  
Qy 481 GARGCTTGCACACATCTTGGGAAGCGGCTGTACCTCTTGGGAAGTGGTTACACCACTC 540  
Db 481 GARGCTTGCACACATCTTGGGAAGCGGCTGTACCTCTTGGGAAGTGGTTACACCACTC 540  
Qy 541 TTGGGAGCTCTGTGACGAAGACCCCGGTACATTTTGGGRACCAACGACATCC 600  
Db 541 TTGGGAGCTCTGTGACGAAGACCCCGGTACATTTTGGGRACCAACGACATCC 600  
Qy 601 MAAGTATGGGAAAAGTTCCTCCCGCAAGACAAAAAGCCCTTAAGACGTATTTCTGARAAT 660  
Db 601 MAAGTATGGGAAAAGTTCCTCCCGCAAGACAAAAAGCCCTTAAGACGTATTTCTGARAAT 660  
Qy 661 TGGGAMCAATTTGACCTCAGACACTAAGAAAAGAACGACTTATATCTTCTGCAAGTGC 720  
Db 661 TGGGAMCAATTTGACCTCAGACACTAAGAAAAGAACGACTTATATCTTCTGCAAGTGC 720  
Qy 721 GCCTGSCACTCTGAGGGAAGTATAAATTAACACCACTTTACAGCTAGACYCTCTTTG 780  
Db 721 GCCTGSCACTCTGAGGGAAGTATAAATTAACACCACTTTACAGCTAGACYCTCTTTG 780  
Qy 781 TAGAAAAGGCAATGAGTGAAGTGCCTAAGTACAACTTTCTTTCAATTAAGAGACAA 840  
Db 781 TAGAAAAGGCAATGAGTGAAGTGCCTAAGTACAACTTTCTTTCAATTAAGAGACAA 840  
Qy 841 CTCACAAATATGTAAAAAGTGATTTATGCCCTACAGGAAGCCCTTACAGTCTACCTCC 900  
Db 841 CTCACAAATATGTAAAAAGTGATTTATGCCCTACAGGAAGCCCTTACAGTCTACCTCC 900  
Qy 901 CTATCCAGCATCCCCGACTCTCTCCCAATTAAGAGACCCCTTCAACCCCAATGG 960  
Db 901 CTATCCAGCATCCCCGACTCTCTCCCAATTAAGAGACCCCTTCAACCCCAATGG 960  
Qy 961 TCCAAAAGGAGATAGACAAAAGGTTAAACAGTGAACCAAGAGTGCCAAATATTCGCCAAT 1020  
Db 961 TCCAAAAGGAGATAGACAAAAGGTTAAACAGTGAACCAAGAGTGCCAAATATTCGCCAAT 1020  
Qy 1021 TATGACCCCTCCAGCAGTGGGAGAGAGATTCGGCCAGCCAGCAGTGCATGCVT 1080  
Db 1021 TATGACCCCTCCAGCAGTGGGAGAGAGATTCGGCCAGCCAGCAGTGCATGCVT 1080  
Qy 1081 TTTTYYTCTCCAGACTTAAAGCAAAATAAAAAAGAGCTTAGGTAAATTTCTCAGATAAYCCT 1140  
Db 1081 TTTTYYTCTCCAGACTTAAAGCAAAATAAAAAAGAGCTTAGGTAAATTTCTCAGATAAYCCT 1140  
Qy 1141 GATGGCTATATTGRTGTTTACAGGGTTAGGACAAATTTCTTGATCTGACATGGAGAT 1200  
Db 1141 GATGGCTATATTGRTGTTTACAGGGTTAGGACAAATTTCTTGATCTGACATGGAGAT 1200  
Qy 1201 ATATATGTCACTCTAAATCAGACACTAACCCCAATGAGAGAGTGCCACCACTAAGTGC 1260  
Db 1201 ATATATGTCACTCTAAATCAGACACTAACCCCAATGAGAGAGTGCCACCACTAAGTGC 1260  
Qy 1261 AGCCTGAGRTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGATGACAAACA 1320  
Db 1261 AGCCTGAGRTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGATGACAAACA 1320  
Qy 1321 GAAGGAAAGANAATGATTTCCCAAGCCAGCARGCAGTTCAGTCTASACCCCTCATTTG 1380  
Db 1321 GAAGGAAAGANAATGATTTCCCAAGCCAGCARGCAGTTCAGTCTASACCCCTCATTTG 1380  
Qy 1381 GGGACACAGAAATCAGTAACTGGGAGATTTGGTGTGCGACACATTTGCTAACTTGTGTCG 1440  
Db 1381 GGGACACAGAAATCAGTAACTGGGAGATTTGGTGTGCGACACATTTGCTAACTTGTGTCG 1440  
Qy 1441 TASAAGGACTAAGGAAAACTTASGAGAAATCTAYGAAATTTACTCAATGATGTCCACCATA 1500  
Db 1441 TASAAGGACTAAGGAAAACTTASGAGAAATCTAYGAAATTTACTCAATGATGTCCACCATA 1500

Qy 1501 ACACAGGGAGGGAAGAAATCCTACTGCGCTTCTGAGAGACTAAGGGAGGCATTGAG 1560  
Db 1501 ACACAGGGAGGGAAGAAATCCTACTGCGCTTCTGAGAGACTAAGGGAGGCATTGAG 1560  
Qy 1561 GAAGCGTGCCTCTCTGTCACTGACTCTTCTGAAAGGCCAACTAATCTTTAAAGCGTAAGTT 1620  
Db 1561 GAAGCGTGCCTCTCTGTCACTGACTCTTCTGAAAGGCCAACTAATCTTTAAAGCGTAAGTT 1620  
Qy 1621 TATCACTCAGTCAGTCAGACATTTAGAAAAAACTTTCAAAAAGTCTGCGGTAGGCCCGGAG 1680  
Db 1621 TATCACTCAGTCAGTCAGACATTTAGAAAAAACTTTCAAAAAGTCTGCGGTAGGCCCGGAG 1680  
Qy 1681 CAAAACCTTAGAAAACTTTGAACTTGGCAACCTTGAATTTTAAATAGAGATCAGGAG 1740  
Db 1681 CAAAACCTTAGAAAACTTTGAACTTGGCAACCTTGAATTTTAAATAGAGATCAGGAG 1740  
Qy 1741 GAGCAGGCGGAACAGGACAAACGGGATTTAAAAAAAGGCCACCGCTTTAGTCATGACCT 1800  
Db 1741 GAGCAGGCGGAACAGGACAAACGGGATTTAAAAAAAGGCCACCGCTTTAGTCATGACCT 1800  
Qy 1801 CAGGCAAGTGAATTTGGAGGCTCTGAAAAAGGAAAAAGCTGGGCAAAATTTGAATGCCCTAA 1860  
Db 1801 CAGGCAAGTGAATTTGGAGGCTCTGAAAAAGGAAAAAGCTGGGCAAAATTTGAATGCCCTAA 1860  
Qy 1861 TAGGCTTGTCTCCAGTCCGCTCTACAGGACACTTTTAAAAAAGATTCTCCAAGTAGAAG 1920  
Db 1861 TAGGCTTGTCTCCAGTCCGCTCTACAGGACACTTTTAAAAAAGATTCTCCAAGTAGAAG 1920  
Qy 1921 TAAGCGGCCCTTCTGTCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCC 1980  
Db 1921 TAAGCGGCCCTTCTGTCATGCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCC 1980  
Qy 1981 CAGGGAACAAAGGCTTTTGTAGTCAGAAAGCACTAAACAGATGATCCAGCAGCAGACTG 2040  
Db 1981 CAGGGAACAAAGGCTTTTGTAGTCAGAAAGCACTAAACAGATGATCCAGCAGCAGACTG 2040  
Qy 2041 AGGCTGCTGGGCAAGGCCCATCCATGCCATCACTCAGAGCCCTGGGTATGCTT 2100  
Db 2041 AGGCTGCTGGGCAAGGCCCATCCATGCCATCACTCAGAGCCCTGGGTATGCTT 2100  
Qy 2101 GACCATTGAGGGCCAGGAAGGTTGCTCTCGACACTGCTGCGGTCTTTAGTCTTACT 2160  
Db 2101 GACCATTGAGGGCCAGGAAGGTTGCTCTCGACACTGCTGCGGTCTTTAGTCTTACT 2160  
Qy 2161 CTTCGTCTCCGGAACAACTGCTCTCAGATCTGTCTAATTTCTGAGGGGGTCCNTAAGAGC 2220  
Db 2161 CTTCGTCTCCGGAACAACTGCTCTCAGATCTGTCTAATTTCTGAGGGGGTCCNTAAGAGC 2220  
Qy 2221 GGCAGTCACTAGATACTTTTCCAGGCCACTAAGTTTATGAATCTGGGAGCTTTTCTTTT 2280  
Db 2221 GGCAGTCACTAGATACTTTTCCAGGCCACTAAGTTTATGAATCTGGGAGCTTTTCTTTT 2280  
Qy 2281 TCACATGCTTTTCTAATTTATGCTTGAAGCCCCCACTAAGCTTTAGGAGAGACATTTCTA 2340  
Db 2281 TCACATGCTTTTCTAATTTATGCTTGAAGCCCCCACTAAGCTTTAGGAGAGACATTTCTA 2340  
Qy 2341 GCAAAAGCAGGGGCCATTTATACCTGAACATAGAGAAAGAACACCCCGTTTGTGTCNC 2400  
Db 2341 GCAAAAGCAGGGGCCATTTATACCTGAACATAGAGAAAGAACACCCCGTTTGTGTCNC 2400  
Qy 2401 CCTGCTTGAAGNAGAAATTAATCTGGAAGTCTGGGCAACAGNAGACATATGGAGCAGC 2460  
Db 2401 CCTGCTTGAAGNAGAAATTAATCTGGAAGTCTGGGCAACAGNAGACATATGGAGCAGC 2460  
Qy 2461 CAAAGATGCCCGTCTGTTCAAGTTAAATTAAGGATTTCCACTTTCCCTTACCAAA 2520  
Db 2461 CAAAGATGCCCGTCTGTTCAAGTTAAATTAAGGATTTCCACTTTCCCTTACCAAA 2520  
Qy 2521 GGCAGTACCCCTCAGACCCAGGCCCAACAGGATTTCAAAAGATTCTTAAGGACTTAA 2580  
Db 2521 GGCAGTACCCCTCAGACCCAGGCCCAACAGGATTTCAAAAGATTCTTAAGGACTTAA 2580

QY 2581 AAGCCCAAGGCTTAGTAAACCATGCAATAACTCCCTGCGACTTAATTCGGTAGTGGATTGAG 2640  
DB 2581 AAGCCCAAGGCTTAGTAAACCATGCAATAACTCCCTGCGACTTAATTCGGTAGTGGATTGAG 2640  
QY 2641 GAGGCAAGAAACCCAGTGGACAGTGGAGGGTTAGTCAAGAGATCTCAGGATTAATCAATGG 2700  
DB 2641 GAGGCAAGAAACCCAGTGGACAGTGGAGGGTTAGTCAAGAGATCTCAGGATTAATCAATGG 2700  
QY 2701 AGGCGGTGTCCTTTATACCCAGCTGTACTAGCCCTTATCTACTGTGMYTTCGAAATAC 2760  
DB 2701 AGGCGGTGTCCTTTATACCCAGCTGTACTAGCCCTTATCTACTGTGMYTTCGAAATAC 2760  
QY 2761 CAGAGGAAGCAGAGTGGTTTACASTCTCTGGACCTTMAAGGATGCTTCTCTGTCATCCCTG 2820  
DB 2761 CAGAGGAAGCAGAGTGGTTTACASTCTCTGGACCTTMAAGGATGCTTCTCTGTCATCCCTG 2820  
QY 2821 TACATCTGACTCTCAATCTCTGTTTGGCTTTGAAGATACTTCAAAACCCARCATCTCAAC 2880  
DB 2821 TACATCTGACTCTCAATCTCTGTTTGGCTTTGAAGATACTTCAAAACCCARCATCTCAAC 2880  
QY 2881 TCACCTGGACTRTTTTACCCCAAGGTTTCAAGGATAGYCCCACTATCTATTTGGCCAGGCAT 2940  
DB 2881 TCACCTGGACTRTTTTACCCCAAGGTTTCAAGGATAGYCCCACTATCTATTTGGCCAGGCAT 2940  
QY 2941 TAGCCCAAGACTTGAGYCAATYMTCACTCTGACACTCTTGTCTCTCRGTAKTGGATG 3000  
DB 2941 TAGCCCAAGACTTGAGYCAATYMTCACTCTGACACTCTTGTCTCTCRGTAKTGGATG 3000  
QY 3001 ATTTACTTTTTRGCGCYRTTTCAGAAACCTTGTGCCATCAAGCCACCAAGRCCTCTTWA 3060  
DB 3001 ATTTACTTTTTRGCGCYRTTTCAGAAACCTTGTGCCATCAAGCCACCAAGRCCTCTTWA 3060  
QY 3061 ATTTCTCTGACCTGTGGCTACAWGGTTTCCAAACSAARAGTCTCARCTCTGCTCACAGC 3120  
DB 3061 ATTTCTCTGACCTGTGGCTACAWGGTTTCCAAACSAARAGTCTCARCTCTGCTCACAGC 3120  
QY 3121 AGGTTAAATCTTAGRCCTAARATTAATCAAGGACCAAGGCCCTCAGTGAGGAAYRYA 3180  
DB 3121 AGGTTAAATCTTAGRCCTAARATTAATCAAGGACCAAGGCCCTCAGTGAGGAAYRYA 3180  
QY 3181 TCCAGGCTTACTTGGCTTATCTCTCATCYCAAAACCTTAAAGCAACTAAGRGRTTCTTGG 3240  
DB 3181 TCCAGGCTTACTTGGCTTATCTCTCATCYCAAAACCTTAAAGCAACTAAGRGRTTCTTGG 3240  
QY 3241 GRTAAAYAGYTTCTCCGAANAATGGAATCCCAGGTTGGCRRAATAGCCAGGYCAITTA 3300  
DB 3241 GRTAAAYAGYTTCTCCGAANAATGGAATCCCAGGTTGGCRRAATAGCCAGGYCAITTA 3300  
QY 3301 WATACASTAATTAAGAAACTCAGAAAGCCATACCCATTTARTAAAGATGGAYAMCTGAA 3360  
DB 3301 WATACASTAATTAAGAAACTCAGAAAGCCATACCCATTTARTAAAGATGGAYAMCTGAA 3360  
QY 3361 GYMRAAGTGGCTTTCCAGGCCCTTAAAGAGGCTTAAACCCAAAGYCCCAAGTGTAAAGYT 3420  
DB 3361 GYMRAAGTGGCTTTCCAGGCCCTTAAAGAGGCTTAAACCCAAAGYCCCAAGTGTAAAGYT 3420  
QY 3421 TGCCAAACRGGGCAAGACTTTTSTTYATYRTCAAGAAAAAACAAGAAAYAGCTCTRGA 3480  
DB 3421 TGCCAAACRGGGCAAGACTTTTSTTYATYRTCAAGAAAAAACAAGAAAYAGCTCTRGA 3480  
QY 3481 GTCTTTACAGACTCCRAGGAYAGCTTGCACACCTTACCCATTTARTAAAGATGGAYAMCTGAA 3540  
DB 3481 GTCTTTACAGACTCCRAGGAYAGCTTGCACACCTTACCCATTTARTAAAGATGGAYAMCTGAA 3540  
QY 3541 GATGTAGTGGCAAGGTTGRCYTCATCTGTTTAYGGGTAGTGGTGGCAGTACAGTYKTA 3600  
DB 3541 GATGTAGTGGCAAGGTTGRCYTCATCTGTTTAYGGGTAGTGGTGGCAGTACAGTYKTA 3600  
QY 3601 GTATCTGAAGCAGTTAAATATAACAGGAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
DB 3601 GTATCTGAAGCAGTTAAATATAACAGGAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
QY 3661 GTGAAYRGCACTCACTGCTAAAGGAGACTTGTGGCTGTGACACAACYGTTTACTTAAA 3720

DB 3661 GTGAAYRGCACTCACTGCTAAAGGAGACTTGTGGCTGTGACACAACYGTTTACTTAAA 3720  
QY 3721 TRTCAGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACTTGTGCAACTCTTAAAC 3780  
DB 3721 TRTCAGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACTTGTGCAACTCTTAAAC 3780  
QY 3781 CCAGYCNCAATTTCTCCAGACAATGAAGAAAGATARAAYATAACTGTCAACAARTAAAT 3840  
DB 3781 CCAGYCNCAATTTCTCCAGACAATGAAGAAAGATARAAYATAACTGTCAACAARTAAAT 3840  
QY 3841 TCTCAAAACCTATGCCACTCGAGGGACCTTGTAGARGTTCCTTGTGACTGATCCYGACCTT 3900  
DB 3841 TCTCAAAACCTATGCCACTCGAGGGACCTTGTAGARGTTCCTTGTGACTGATCCYGACCTT 3900  
QY 3901 CAACTTGTATCTAGTGAAGTTCCTTGTAGAAAAAGGACTTCGAAAAAGYGGGGTATGC 3960  
DB 3901 CAACTTGTATCTAGTGAAGTTCCTTGTAGAAAAAGGACTTCGAAAAAGYGGGGTATGC 3960  
QY 3961 AGTGGTCACTGATTAATGGAATATTTGAAAGTAAATCCCTCACTCCAGGAACTAGTCTYA 4020  
DB 3961 AGTGGTCACTGATTAATGGAATATTTGAAAGTAAATCCCTCACTCCAGGAACTAGTCTYA 4020  
QY 4021 GCTRCAGAACTTAATAGCCCTCACTYTKGGGCACTAGAAATTTAGGAGAGRAAAAAAGGYAAA 4080  
DB 4021 GCTRCAGAACTTAATAGCCCTCACTYTKGGGCACTAGAAATTTAGGAGAGRAAAAAAGGYAAA 4080  
QY 4081 TATATATACAGACTCTTRARTATGCTVACCTAGTCTNCTCAATGCCATGMRGCAATATGSAR 4140  
DB 4081 TATATATACAGACTCTTRARTATGCTVACCTAGTCTNCTCAATGCCATGMRGCAATATGSAR 4140  
QY 4141 AGAAAGGGAATTCCTAACTTCYAGRGAAACCTTATCAACATCAGAAAGCCATTAGGAR 4200  
DB 4141 AGAAAGGGAATTCCTAACTTCYAGRGAAACCTTATCAACATCAGAAAGCCATTAGGAR 4200  
QY 4201 ATTTATTTAYTGGCGWTACAGAAACCTTARAGAGTGGHAGTCTTACATGCTGCGGGTCACTCA 4260  
DB 4201 ATTTATTTAYTGGCGWTACAGAAACCTTARAGAGTGGHAGTCTTACATGCTGCGGGTCACTCA 4260  
QY 4261 NAAAGAAAGRAAGGAATASAGRGAAATYTGCCAAAGCAKATATTTGAAGCAAAAGAGC 4320  
DB 4261 NAAAGAAAGRAAGGAATASAGRGAAATYTGCCAAAGCAKATATTTGAAGCAAAAGAGC 4320  
QY 4321 TGCAAGGCAAGGACCTTCACTTAAAGTGTATTAACCTTCCCTTAGTATAGGTTAATCC 4380  
DB 4321 TGCAAGGCAAGGACCTTCACTTAAAGTGTATTAACCTTCCCTTAGTATAGGTTAATCC 4380  
QY 4381 CTTCCGGGAAACCAAGCCCTCAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCAGGAG 4440  
DB 4381 CTTCCGGGAAACCAAGCCCTCAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCAGGAG 4440  
QY 4441 CAGTTTCTCCCTCGGGACGTTAGCCACTGAAGAGGGAATACTTTTGCCTGCAAC 4500  
DB 4441 CAGTTTCTCCCTCGGGACGTTAGCCACTGAAGAGGGAATACTTTTGCCTGCAAC 4500  
QY 4501 TATCAATGGAAATTAATTAACCTTCAATCAAACTTCTAGGCACTAGATAGCAC 4560  
DB 4501 TATCAATGGAAATTAATTAACCTTCAATCAAACTTCTAGGCACTAGATAGCAC 4560  
QY 4561 CCATCARATGGCCAAATCAATTAATTTACTGGAACGAGCCCTTTTCAAACTATCAAGCARAT 4620  
DB 4561 CCATCARATGGCCAAATCAATTAATTTACTGGAACGAGCCCTTTTCAAACTATCAAGCARAT 4620  
QY 4621 AKTCAAGGCTGTGAAKTGTGCCARAAATTAATCCCTGCTCTATCGCAGAGCTCTTC 4680  
DB 4621 AKTCAAGGCTGTGAAKTGTGCCARAAATTAATCCCTGCTCTATCGCAGAGCTCTTC 4680  
QY 4681 AGGABAAACAAARACACAGGCCATTTACCTGRRABAACTGGCAACTGATTTTACCCACAAG 4740  
DB 4681 AGGABAAACAAARACACAGGCCATTTACCTGRRABAACTGGCAACTGATTTTACCCACAAG 4740  
QY 4741 CCCAAACCTCAGGGAATTCAGTATCTACTAGTCTGGGTARATATCTTTCACGSGTTGGCA 4800

Db 4741 CCCAAACCTCAGGGATTTTCAGTATCTACTAGTCTGGGTARATACATTTTCAGGGGTTGGCA 4800  
Qy  
4801 RAGGCTTTCCCTGTAGGACAGAAAAGCCGAGAGGTAAATAAGGCATAGTTTCATGAA 4860  
Db  
4801 RAGGCTTTCCCTGTAGGACAGAAAAGCCGAGAGGTAAATAAGGCATAGTTTCATGAA 4860  
Qy  
4861 ATAATTTCCAGATTCCGAGCTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAG 4920  
Db  
4861 ATAATTTCCAGATTCCGAGCTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAG 4920  
Qy  
4921 GCCACAGTAACCCAGGGAGTATCCAGGCGTTAGGTATACGATATCATCTACACTGCGCC 4980  
Db  
4921 GCCACAGTAACCCAGGGAGTATCCAGGCGTTAGGTATACGATATCATCTACACTGCGCC 4980  
Qy  
4981 TGAAGGCCACAGTCTCTAGGGAAGTTCAGAGAAATGAATGAAYAACTCAAAGGACATCTA 5040  
Db  
4981 TGAAGGCCACAGTCTCTAGGGAAGTTCAGAGAAATGAATGAAYAACTCAAAGGACATCTA 5040  
Qy  
5041 AAAAAAGCAACCCAGGAAACCCACCTCACATGGCTGVTCTGTTGCCCTATAGCCCTTAAAA 5100  
Db  
5041 AAAAAAGCAACCCAGGAAACCCACCTCACATGGCTGVTCTGTTGCCCTATAGCCCTTAAAA 5100  
Qy  
5101 AGAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATAGAAATGCTGTATGGAAGG 5160  
Db  
5101 AGAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATAGAAATGCTGTATGGAAGG 5160  
Qy  
5161 CCCTTCATAACCAATGACTTGTGCTTGACCCOAGACAGCCAACTTAGTTGACAGATCA 5220  
Db  
5161 CCCTTCATAACCAATGACTTGTGCTTGACCCOAGACAGCCAACTTAGTTGACAGATCA 5220  
Qy  
5221 CCTCCTTAGCCAAATATCAACAGTTCTTAAACATTAACAAGAACCTTATCCCTGAGAAG 5280  
Db  
5221 CCTCCTTAGCCAAATATCAACAGTTCTTAAACATTAACAAGAACCTTATCCCTGAGAAG 5280  
Qy  
5281 AGGGAAAAGAACTATTCCACCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATT 5340  
Db  
5281 AGGGAAAAGAACTATTCCACCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATT 5340  
Qy  
5341 CCCCATCCCTAGATACATCTCTGGGAAGGACCCCTACCCAGTCAATTTATTTATTTACCCCACTG 5400  
Db  
5341 CCCCATCCCTAGATACATCTCTGGGAAGGACCCCTACCCAGTCAATTTATTTATTTACCCCACTG 5400  
Qy  
5401 CGGTTAAAGTGGCTGAGTGGAGTCTTGGATACATCACATTTGAGTCAAAATCCTGGATAC 5460  
Db  
5401 CGGTTAAAGTGGCTGAGTGGAGTCTTGGATACATCACATTTGAGTCAAAATCCTGGATAC 5460  
Qy  
5461 TGCCAAAGGAACTCGAAATCCAGGAGACAAAGCTAGCTATTTCCTGTGAACCTCTAGAGG 5520  
Db  
5461 TGCCAAAGGAACTCGAAATCCAGGAGACAAAGCTAGCTATTTCCTGTGAACCTCTAGAGG 5520  
Qy  
5521 ATTTGGGCTGCTCTTCAAAACAAACCCAGGAGGAAAGTAACTAAATCATAAATCCGCC 5580  
Db  
5521 ATTTGGGCTGCTCTTCAAAACAAACCCAGGAGGAAAGTAACTAAATCATAAATCCGCC 5580  
Qy  
5581 ATGSGCCCTCCCTATCATATTTTCTCTKTASTGTTSTTTTACCTSTTTCACTCTCACT 5640  
Db  
5581 ATGSGCCCTCCCTATCATATTTTCTCTKTASTGTTSTTTTACCTSTTTCACTCTCACT 5640  
Qy  
5641 GCACCCCTCCATGCGCTGTATGACCCAGTACTCCCCCTYACCMAGAGTTTCTATGGAGA 5700  
Db  
5641 GCACCCCTCCATGCGCTGTATGACCCAGTACTCCCCCTYACCMAGAGTTTCTATGGAGA 5700  
Qy  
5701 ATGACAGGTCGCGGAAATATGATGCCCCCATCGTATAGAGTCTTTTSTAAGGAAACCCCC 5760  
Db  
5701 ATGACAGGTCGCGGAAATATGATGCCCCCATCGTATAGAGTCTTTTSTAAGGAAACCCCC 5760  
Qy  
5761 ACCTTCAGTCCCAACCCATATGCCCCGCACTGCTATCACTCTGCGCACTCTTTGCAATG 5820  
Db  
5761 ACCTTCAGTCCCAACCCATATGCCCCGCACTGCTATCACTCTGCGCACTCTTTGCAATG 5820  
Qy  
5821 CATGAAATACCTATTATTTGACAGGAAATGATTAATCTCTAGTTGCTCTGGAGACTT 5880  
Db  
5821 CATGAAATACCTATTATTTGACAGGAAATGATTAATCTCTAGTTGCTCTGGAGACTT 5880

Qy 5881 GGAGTCACTGCTGTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGTGGAGTT 5940  
Db 5881 GGAGTCACTGCTGTGGACTTACTTCACCCAAACTGGTATGTCTGATGGGGTGGAGTT 5940  
Qy 5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAAGAAATCTCCCAACTCACCGGGGTACAT 6000  
Db 5941 CAAGATCAGGCAAGAGAAAAACATGTAAAAAGAAATCTCCCAACTCACCGGGGTACAT 6000  
Qy 6001 GGCACTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTCATATGAACCCCTCCGTACC 6060  
Db 6001 GGCACTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTCATATGAACCCCTCCGTACC 6060  
Qy 6061 CATACTCGCTGTAAAGCTATTAAATACCAACCTCTACTGGGCTCCATGAGGTCTCGGCC 6120  
Db 6061 CATACTCGCTGTAAAGCTATTAAATACCAACCTCTACTGGGCTCCATGAGGTCTCGGCC 6120  
Qy 6121 CAAAACCTACTAACTGTTGGATATGCTCCCTCCCTGAACCTTCARGCCATATGTTTCAATC 6180  
Db 6121 CAAAACCTACTAACTGTTGGATATGCTCCCTCCCTGAACCTTCARGCCATATGTTTCAATC 6180  
Qy 6181 CCTGTACTGAACAAATGGAAACAACTTCAGCAAGAAATAAACACACCTTCCTGGTTTGTGA 6240  
Db 6181 CCTGTACTGAACAAATGGAAACAACTTCAGCAAGAAATAAACACACCTTCCTGGTTTGTGA 6240  
Qy 6241 GGACCTCTTGTTCCTCAATSTGGAATTAACCCATACCTCAACCTCAACCTGTTGAAAAATTT 6300  
Db 6241 GGACCTCTTGTTCCTCAATSTGGAATTAACCCATACCTCAACCTCAACCTGTTGAAAAATTT 6300  
Qy 6301 AGCAATACTACATACACAAACCACTCCCAATGCAATCAGGTGGGTAACTCTCCCAACACAA 6360  
Db 6301 AGCAATACTACATACACAAACCACTCCCAATGCAATCAGGTGGGTAACTCTCTCCCAACAA 6360  
Qy 6361 ATAGTCTGCTACCTCAGGAATAATTTTGTCTGTGGTACTCAGCCTATGCTGTTGTTG 6420  
Db 6361 ATAGTCTGCTACCTCAGGAATAATTTTGTCTGTGGTACTCAGCCTATGCTGTTGTTG 6420  
Qy 6421 AATGGCTCTCAGAACTATGCTGCTCTCATCTTCTAGTGGCCCTCATGRCATCTAC 6480  
Db 6421 AATGGCTCTCAGAACTATGCTGCTCTCATCTTCTAGTGGCCCTCATGRCATCTAC 6480  
Qy 6481 ACTGAACAAGATTTATACAGTTATGTCTATATCTAAGCCCGCAACAAAGAGTACCCATT 6540  
Db 6481 ACTGAACAAGATTTATACAGTTATGTCTATATCTAAGCCCGCAACAAAGAGTACCCATT 6540  
Qy 6541 CTTCTTTTGTATAGGACAGGAGTGTAGGTGCACTAGGTACTGGCACTTGGCGGTATC 6600  
Db 6541 CTTCTTTTGTATAGGACAGGAGTGTAGGTGCACTAGGTACTGGCACTTGGCGGTATC 6600  
Qy 6601 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660  
Db 6601 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660  
Qy 6661 GTGCGGACTCCCTGGTCACTTGAAGTCACTTAACTCCCTAGCAGCAGTAGTCTCT 6720  
Db 6661 GTGCGGACTCCCTGGTCACTTGAAGTCACTTAACTCCCTAGCAGCAGTAGTCTCT 6720  
Qy 6721 CRAAATCGAAGACTTTAGACTYGTAAACCCCTGAGAGAGGGGAACTGTGTTATTTTTTA 6780  
Db 6721 CRAAATCGAAGACTTTAGACTYGTAAACCCCTGAGAGAGGGGAACTGTGTTATTTTTTA 6780  
Qy 6781 GGGGAGAACTGCTGTTATTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAA 6840  
Db 6781 GGGGAGAACTGCTGTTATTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAA 6840  
Qy 6841 ATTSAGATCGAATAACACGTAKAGCAGARAGCTTCGAAACACTGGACCCCTGGGCGCTC 6900  
Db 6841 ATTSAGATCGAATAACACGTAKAGCAGARAGCTTCGAAACACTGGACCCCTGGGCGCTC 6900  
Qy 6901 CTCAGCCATGATGCCCCCTGATTTCTCCCTTTTAGGACCTCTAGCAGCTATAATATTG 6960  
Db 6901 CTCAGCCATGATGCCCCCTGATTTCTCCCTTTTAGGACCTCTAGCAGCTATAATATTG 6960





Db 1081 TTTYVTCTCCAGACTTAAGCAAAATAAAAAACAGACTTAGGTAAATTTCTCAGATAAYCCT 1140  
Qy 1141 GATGCTATATGRTGTTTTACAAGGGTTAGGACAAATTTCTTTGATCTGACATGGAGAGAT 1200  
Db 1141 GATGCTATATGRTGTTTTACAAGGGTTAGGACAAATTTCTTTGATCTGACATGGAGAGAT 1200  
Qy 1201 ATATATGTCACCTGCTAAATCAGACACTAACCCCAATGAGAGAAGTGCACCAATACCTGC 1260  
Db 1201 ATATATGTCACCTGCTAAATCAGACACTAACCCCAATGAGAGAAGTGCACCAATACCTGC 1260  
Qy 1261 AGCTGAGRTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGATGANGATGACAACA 1320  
Db 1261 AGCTGAGRTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGATGANGATGACAACA 1320  
Qy 1321 GAAGAAAGANAATGATTTCCCAAGGCCAGCARGAGTTCCAGTCTASACCCCTCATTTG 1380  
Db 1321 GAAGAAAGANAATGATTTCCCAAGGCCAGCARGAGTTCCAGTCTASACCCCTCATTTG 1380  
Qy 1381 GGGACACAGAAATCAGTAACATGGGAGATTGGTGTGCAGACATTTGCTTAAGTGTGTGC 1440  
Db 1381 GGGACACAGAAATCAGTAACATGGGAGATTGGTGTGCAGACATTTGCTTAAGTGTGTGC 1440  
Qy 1441 TASAGGACTAAGGAAAACCTASGAAGAAATCTAYGAATTTACTCAATGATGTCCACCATA 1500  
Db 1441 TASAGGACTAAGGAAAACCTASGAAGAAATCTAYGAATTTACTCAATGATGTCCACCATA 1500  
Qy 1501 ACACAGGGAGAGGAGAAAATCTTACTGCTTTCTGGAGAGACTAAAGGAGGCATTTGAG 1560  
Db 1501 ACACAGGGAGAGGAGAAAATCTTACTGCTTTCTGGAGAGACTAAAGGAGGCATTTGAG 1560  
Qy 1561 GAAGCGTGCTCTCTGTCACTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTT 1620  
Db 1561 GAAGCGTGCTCTCTGTCACTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTT 1620  
Qy 1621 TATCACTCAGTCAGTCAGACATTTAGAAAAAATCTTCAAAAGTCTGCGGTAGGCCCGGAG 1680  
Db 1621 TATCACTCAGTCAGTCAGACATTTAGAAAAAATCTTCAAAAGTCTGCGGTAGGCCCGGAG 1680  
Qy 1681 CAAAACCTTAGAAACCTATTGAACCTTGGCAACCTGCGTTTTTTTAAATAGAGATCAGGAG 1740  
Db 1681 CAAAACCTTAGAAACCTATTGAACCTTGGCAACCTGCGTTTTTTTAAATAGAGATCAGGAG 1740  
Qy 1741 GAGCAGGCGGAACAGGACAAACGGGATTTAAABAAAGGCCACCGCTTTAGTCATCACCT 1800  
Db 1741 GAGCAGGCGGAACAGGACAAACGGGATTTAAABAAAGGCCACCGCTTTAGTCATCACCT 1800  
Qy 1801 CAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGAAAAGCTGGGCAAAATTTGAATGCCTAA 1860  
Db 1801 CAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGAAAAGCTGGGCAAAATTTGAATGCCTAA 1860  
Qy 1861 TAGGGCTTGCTTTCCAGTGGGCTTACAAGGACACTTTTAAAAAGATTTGTCCAAGTAGAAG 1920  
Db 1861 TAGGGCTTGCTTTCCAGTGGGCTTACAAGGACACTTTTAAAAAGATTTGTCCAAGTAGAAG 1920  
Qy 1921 TAAGCGCCCTCTGTCATGCGCTTATTTCAAGGATCACTGGAGGCCCATCTGCC 1980  
Db 1921 TAAGCGCCCTCTGTCATGCGCTTATTTCAAGGATCACTGGAGGCCCATCTGCC 1980  
Qy 1981 CAGGGACAAAGGTCTTTTGAAGTCAGAACCTAACAGATGATCCAGCAGCAGACTG 2040  
Db 1981 CAGGGACAAAGGTCTTTTGAAGTCAGAACCTAACAGATGATCCAGCAGCAGACTG 2040  
Qy 2041 AGGGTCTTGGGCAAGGCCATCCCATGCCATCACTCAAGAGCCCTGGGTATGCTT 2100  
Db 2041 AGGGTCTTGGGCAAGGCCATCCCATGCCATCACTCAAGAGCCCTGGGTATGCTT 2100  
Qy 2101 GACCATGAGGCGCAGGAGGTGCTCTCTGACACTGGTGGCTTCTTAGTCTTACT 2160  
Db 2101 GACCATGAGGCGCAGGAGGTGCTCTCTGACACTGGTGGCTTCTTAGTCTTACT 2160  
Qy 2161 CTTCTGTCGGGACAACTGCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGCG 2220  
Db 2161 CTTCTGTCGGGACAACTGCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGCG 2220

Qy 2221 GGCAGTCACTAGATACCTTTTCCAGCCACTAAAGTTATGAACCTGGGAGCTTTATCTTTT 2280  
Db 2221 GGCAGTCACTAGATACCTTTTCCAGCCACTAAAGTTATGAACCTGGGAGCTTTATCTTTT 2280  
Qy 2281 TCACATGCTTTTCTAAATTTATGTTGAAAGCCCACTACCTTTGTTAGGAGAGACATTTCTA 2340  
Db 2281 TCACATGCTTTTCTAAATTTATGTTGAAAGCCCACTACCTTTGTTAGGAGAGACATTTCTA 2340  
Qy 2341 GCAAAGCAGGGCCATTATACCTGAAATAGGAGAAAGAACCCCGTTTGTGTGCC 2400  
Db 2341 GCAAAGCAGGGCCATTATACCTGAAATAGGAGAAAGAACCCCGTTTGTGTGCC 2400  
Qy 2401 CCTGTTGAGAAAGGAATTAATCTGAAAGTCTGGCAACAGAAAGACAAATATGGACGAGC 2460  
Db 2401 CCTGTTGAGAAAGGAATTAATCTGAAAGTCTGGCAACAGAAAGACAAATATGGACGAGC 2460  
Qy 2461 CAAAGAAATGCCCGCTCTGTTCAAGTTAAATTAAGGATTTCACTTTCTCTACCAAA 2520  
Db 2461 CAAAGAAATGCCCGCTCTGTTCAAGTTAAATTAAGGATTTCACTTTCTCTACCAAA 2520  
Qy 2521 GGCAGTACCCCTCAGACCCCAAGGCCCAAGGATTTCCAAAGATTTGTTAAGGACTTAA 2580  
Db 2521 GGCAGTACCCCTCAGACCCCAAGGCCCAAGGATTTCCAAAGATTTGTTAAGGACTTAA 2580  
Qy 2581 AAGCCCAAGGCTTAGTAAACCATGCACTCTCCCTGCAGTAATTTCCGTAGTGGATTGAG 2640  
Db 2581 AAGCCCAAGGCTTAGTAAACCATGCACTCTCCCTGCAGTAATTTCCGTAGTGGATTGAG 2640  
Qy 2641 GAGGCACAGAAACCCAGTGGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTTATCAATGG 2700  
Db 2641 GAGGCACAGAAACCCAGTGGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTTATCAATGG 2700  
Qy 2701 AGGCGTTGTCTTTTATACCCAGCTGPACTAGCCCTTATCTAGCTGMYTTCCCAAAATAC 2760  
Db 2701 AGGCGTTGTCTTTTATACCCAGCTGPACTAGCCCTTATCTAGCTGMYTTCCCAAAATAC 2760  
Qy 2761 CAGAGGAACAGAGTGGTTTACSTCTCGGACCTTMAAGATGCTTCTTCTGCATCCCTG 2820  
Db 2761 CAGAGGAACAGAGTGGTTTACSTCTCGGACCTTMAAGATGCTTCTTCTGCATCCCTG 2820  
Qy 2821 TACATCTGACTCTCAATTTCTTTGCTTTGAAGTACTTTCAAAACCCARCATCTCAAC 2880  
Db 2821 TACATCTGACTCTCAATTTCTTTGCTTTGAAGTACTTTCAAAACCCARCATCTCAAC 2880  
Qy 2881 TCACCTGGACTTTTACCCCAAGGGTTACGGGATAGYCCCCATCTATTTGGCCAGGCAT 2940  
Db 2881 TCACCTGGACTTTTACCCCAAGGGTTACGGGATAGYCCCCATCTATTTGGCCAGGCAT 2940  
Qy 2941 TAGCCCAAGACTTGAGYCACTYMTCACTGGACACTCTTGTCTCTCRGTAKGTGGATG 3000  
Db 2941 TAGCCCAAGACTTGAGYCACTYMTCACTGGACACTCTTGTCTCTCRGTAKGTGGATG 3000  
Qy 3001 ATTTACTTTTGGCYGCCYRTTTCAGAAACCTTTGTGCATCAAGCCACCAAGCRCTCTTMA 3060  
Db 3001 ATTTACTTTTGGCYGCCYRTTTCAGAAACCTTTGTGCATCAAGCCACCAAGCRCTCTTMA 3060  
Qy 3061 ATTTCTCTCGYACCTGTGGCTACAWGGTTTCCAAACSAARAGCTCARCTCTGCTCACAGC 3120  
Db 3061 ATTTCTCTCGYACCTGTGGCTACAWGGTTTCCAAACSAARAGCTCARCTCTGCTCACAGC 3120  
Qy 3121 AGGTTAAATACTTAGGCTTAARATTTTCAAAAGGCCACCGCCCTCAGTGGAGAAATYA 3180  
Db 3121 AGGTTAAATACTTAGGCTTAARATTTTCAAAAGGCCACCGCCCTCAGTGGAGAAATYA 3180  
Qy 3181 TCCAGCTTACTTGGCTTATCTCATCYCAAAAACCTTAAAGCAACTAAGRGRRTTCTTTG 3240  
Db 3181 TCCAGCTTACTTGGCTTATCTCATCYCAAAAACCTTAAAGCAACTAAGRGRRTTCTTTG 3240  
Qy 3241 GCRTAAYAGGYTTCTGCCGAATAAGATTTCCCGAGTWTGGCRAAATAGCCAGGYCATTTA 3300  
Db 3241 GCRTAAYAGGYTTCTGCCGAATAAGATTTCCCGAGTWTGGCRAAATAGCCAGGYCATTTA 3300



3301 QY WATACASTAATTAAAGGAACTCAGAAAGCCAATAACCCATTARTAAAGATGGAAMCTGAA 3360  
3301 Db WATACASTAATTAAAGGAACTCAGAAAGCCAATAACCCATTARTAAAGATGGAAMCTGAA 3360  
3361 QY GYNRAAGTGGCTTTCCAGGGCCCTTAAAGAGGCCCTTAAACCCCAAGYCCCAAGTGTAAAGYT 3420  
3361 Db GYNRAAGTGGCTTTCCAGGGCCCTTAAAGAGGCCCTTAAACCCCAAGYCCCAAGTGTAAAGYT 3420  
3421 QY TGCCACRGGGCAAGACTTTTSTTYATYRTACAGAAAAAACAAGAAAGCTCTRGGG 3480  
3421 Db TGCCACRGGGCAAGACTTTTSTTYATYRTACAGAAAAAACAAGAAAGCTCTRGGG 3480  
3481 QY GTCCTTACACAGRTCCRAGGAYGAGCTGCAACCYRTGCRVACCTGASTAAGGAAAYT 3540  
3481 Db GTCCTTACACAGRTCCRAGGAYGAGCTGCAACCYRTGCRVACCTGASTAAGGAAAYT 3540  
3541 QY GATGTAGTGCAAGGGTTGRCVTCATTGTTTAYGGGTAGTGGCAGTAGCAGTIVKTA 3600  
3541 Db GATGTAGTGCAAGGGTTGRCVTCATTGTTTAYGGGTAGTGGCAGTAGCAGTIVKTA 3600  
3601 QY GTATCTGACGAGCTTAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
3601 Db GTATCTGACGAGCTTAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
3661 QY GTGAAVRGCATCTCACTGCTAAAGGAGACTGTGGCTGTGACACACACGTTTACTTAA 3720  
3661 Db GTGAAVRGCATCTCACTGCTAAAGGAGACTGTGGCTGTGACACACACGTTTACTTAA 3720  
3721 QY TRTCAGGCTTATTAATGGAAGGCCAGTGTGCTGCRACCTGTGCACTTGTGCAACTTTAAC 3780  
3721 Db TRTCAGGCTTATTAATGGAAGGCCAGTGTGCTGCRACCTGTGCAACTTTAAC 3780  
3781 QY CAGYCNCACTTCTCCAGACAATGAAGAAAGATARAAYATAACTGTCAACAARTAAAT 3840  
3781 Db CAGYCNCACTTCTCCAGACAATGAAGAAAGATARAAYATAACTGTCAACAARTAAAT 3840  
3841 QY TCTCAAACTTATCCCACTCGAGGGACCTTGTAGAGTTCCTTGTGACTGATCCYAGCCTT 3900  
3841 Db TCTCAAACTTATCCCACTCGAGGGACCTTGTAGAGTTCCTTGTGACTGATCCYAGCCTT 3900  
3901 QY CAACTTGTATGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGC 3960  
3901 Db CAACTTGTATGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGC 3960  
3961 QY AGTGTCAGTGATAATGGAATAYTTGAAAGTAATCCCTCACTCCAGGAACCTAGTGCTYA 4020  
3961 Db AGTGTCAGTGATAATGGAATAYTTGAAAGTAATCCCTCACTCCAGGAACCTAGTGCTYA 4020  
4021 QY GCTRGCAAGTAATAGCCCTCACTGAGGCACTAGAAATAGGAGAAAGRAAAAAGGYAAA 4080  
4021 Db GCTRGCAAGTAATAGCCCTCACTGAGGCACTAGAAATAGGAGAAAGRAAAAAGGYAAA 4080  
4081 QY TATATATACAGACTCTRARTATGCTYACCTAGTTCNTCCATGCCATGMRGCAATATGSAR 4140  
4081 Db TATATATACAGACTCTRARTATGCTYACCTAGTTCNTCCATGCCATGMRGCAATATGSAR 4140  
4141 QY AGAAGGGAAATTCCTAACTTCYAGRGAACACCTATCAMACATCAGGAAGCCATTAGGAR 4200  
4141 Db AGAAGGGAAATTCCTAACTTCYAGRGAACACCTATCAMACATCAGGAAGCCATTAGGAR 4200  
4201 QY ATTATTATYGGWGTACAGAACTTARAGAGTGGMAGTCTTACATCTCYGGGGTCAATCA 4260  
4201 Db ATTATTATYGGWGTACAGAACTTARAGAGTGGMAGTCTTACATCTCYGGGGTCAATCA 4260  
4261 QY NAAAGGAAGRAAGGGAATASAGRAAYTGCCCAAGCAKATATTTGAAGCHAAAGAGC 4320  
4261 Db NAAAGGAAGRAAGGGAATASAGRAAYTGCCCAAGCAKATATTTGAAGCHAAAGAGC 4320  
4321 QY TGCAAGGCAAGGACCTTCCATTTAGAAATGCTTTATTAACCTTCCCTTAGTATAGGGTAATCC 4380  
4321 Db TGCAAGGCAAGGACCTTCCATTTAGAAATGCTTTATTAACCTTCCCTTAGTATAGGGTAATCC 4380  
4381 QY CTTCCGGGAAACCAAGCCCAAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCACGAGG 4440

4381 Db CTTCCGGGAAACCAAGCCCAAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCACGAGG 4440  
4441 QY CAGTTTTCTCCCTCTCGGACCGTGTAGCCACTGAAGAGGAGAAATACTTTTCTGCTGCAAC 4500  
4441 Db CAGTTTTCTCCCTCTCGGACCGTGTAGCCACTGAAGAGGAGAAATACTTTTCTGCTGCAAC 4500  
4501 QY TATCCAATGGAATATTAATTAACCCCTTCATCAAACTTTCACTTAGGCATCGATAGCAC 4560  
4501 Db TATCCAATGGAATATTAATTAACCCCTTCATCAAACTTTCACTTAGGCATCGATAGCAC 4560  
4561 QY CCATCARATGGCCCAATCATTTATTTACTGACCAAGCCCTTTCAAACTATCAAGCARAT 4620  
4561 Db CCATCARATGGCCCAATCATTTATTTACTGACCAAGCCCTTTCAAACTATCAAGCARAT 4620  
4621 QY AKTCAGGGCTGTGAARTGTGCCARARAAATAATCCCTGCTCTCTATCGCCAAAGCTCTTC 4680  
4621 Db AKTCAGGGCTGTGAARTGTGCCARARAAATAATCCCTGCTCTCTATCGCCAAAGCTCTTC 4680  
4681 QY AGGARAACAAARAACAGGCCATTACCTGRRARAARACTGGCAACTGATTTTACCCACAAG 4740  
4681 Db AGGARAACAAARAACAGGCCATTACCTGRRARAARACTGGCAACTGATTTTACCCACAAG 4740  
4741 QY CCCAAACCTCAGGGATTTTCAAGTATCTACTAGTCTGGGTARATATCTTTCAAGGGTTGGGCA 4800  
4741 Db CCCAAACCTCAGGGATTTTCAAGTATCTACTAGTCTGGGTARATATCTTTCAAGGGTTGGGCA 4800  
4801 QY RAGGCCCTCCCTCTGAGGACAGAAAAAGGCCCAAGAGGTAAATAAGGCACACTAGTTTCATGAA 4860  
4801 Db RAGGCCCTCCCTCTGAGGACAGAAAAAGGCCCAAGAGGTAAATAAGGCACACTAGTTTCATGAA 4860  
4861 QY ATAAATCCAGATTCGGAATTCCTCCGAGGCTTACAGAGTGAACATAGCCCTGCTTTCCAG 4920  
4861 Db ATAAATCCAGATTCGGAATTCCTCCGAGGCTTACAGAGTGAACATAGCCCTGCTTTCCAG 4920  
4921 QY GCCCAGATTAACCCAGGGAGTATCCAGGGCTTAGGTATAGGATATCAGTATCAGTCTGCGCC 4980  
4921 Db GCCCAGATTAACCCAGGGAGTATCCAGGGCTTAGGTATAGGATATCAGTATCAGTCTGCGCC 4980  
4981 QY TGAAGGCCACAGTCCCTCAGGGAAGGTGAGAAATGAATGAATAAGTAAAGGACATCTA 5040  
4981 Db TGAAGGCCACAGTCCCTCAGGGAAGGTGAGAAATGAATGAATAAGTAAAGGACATCTA 5040  
5041 QY AAAAAGCAAAACCCAGGAAACCCCACTCAGATGCGCTGCTGTGTGCTATAGCCTTAAAA 5100  
5041 Db AAAAAGCAAAACCCAGGAAACCCCACTCAGATGCGCTGCTGTGTGCTATAGCCTTAAAA 5100  
5101 QY AGAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGG 5160  
5101 Db AGAATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGG 5160  
5161 QY CCTTTCATTAACCAATGACCTTGTGCTTGACCCCAAGCAGCCACTTAGTTGAGACATCA 5220  
5161 Db CCTTTCATTAACCAATGACCTTGTGCTTGACCCCAAGCAGCCACTTAGTTGAGACATCA 5220  
5221 QY CCTCTTACCAAAATATCAAAAGTCTTAAAAACATTAAAGGAACCTATCCCTGAGAAG 5280  
5221 Db CCTCTTACCAAAATATCAAAAGTCTTAAAAACATTAAAGGAACCTATCCCTGAGAAG 5280  
5281 QY AGGAAAAAGAACTATTTCCACCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAAAT 5340  
5281 Db AGGAAAAAGAACTATTTCCACCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAAAT 5340  
5341 QY CCCCATCCCTAGATACATCTCTGGGAGGACCTTACCCAGTCAATTTATTTACCCCACTG 5400  
5341 Db CCCCATCCCTAGATACATCTCTGGGAGGACCTTACCCAGTCAATTTATTTATTTACCCCACTG 5400  
5401 QY CGGTTAAAGTGTCTGAGTGGAGTCTTGGATACATCACACTTTGAGTCAAACTCTGATAC 5460  
5401 Db CGGTTAAAGTGTCTGAGTGGAGTCTTGGATACATCACACTTTGAGTCAAACTCTGATAC 5460  
5461 QY TGCCAAAGGAACTGAAAAATCCAGGAGCAACCGCTAGCTTATTCCTGTGAACTCTAGAGG 5520  
5461 Db TGCCAAAGGAACTGAAAAATCCAGGAGCAACCGCTAGCTTATTCCTGTGAACTCTAGAGG 5520

Db 5461 TGCCAAAGGAACCTGAAATCCAGGAGACAAACGCTAGCTATTCTCTGTGAAACCTCTAGAGG 5520  
Qy 5521 ATTTGGGCTGCTCTTCAACAAACACAGAGGAAAGTAACATAAATCATAAATCCCC 5580  
Db 5521 ATTTGGGCTGCTCTTCAACAAACACAGAGGAAAGTAACATAAATCATAAATCCCC 5580  
Qy 5581 ATGSGCTCCCTTATCAVATTTTCTCTKTAGTGTSTTTTACCTSTTTTCACTCTCACT 5640  
Db 5581 ATGSGCTCCCTTATCAVATTTTCTCTKTAGTGTSTTTTACCTSTTTTCACTCTCACT 5640  
Qy 5641 GCACCCCTCCATGCGCTGTATGACCACTAGCTCCCTYACCMAGAGTTTCTATGGAGA 5700  
Db 5641 GCACCCCTCCATGCGCTGTATGACCACTAGCTCCCTYACCMAGAGTTTCTATGGAGA 5700  
Qy 5701 ATGACGCTCCGGAAATATTGATGCCCATCGTATAGAGTCTTTTAAAGGAACCC 5760  
Db 5701 ATGACGCTCCGGAAATATTGATGCCCATCGTATAGAGTCTTTTAAAGGAACCC 5760  
Qy 5761 ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCCATG 5820  
Db 5761 ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCCATG 5820  
Qy 5821 CATGCAAAATACCTATTATTTGGACAGGAAAAATGATTAATCTCTAGTTGCTCGAGGACTT 5880  
Db 5821 CATGCAAAATACCTATTATTTGGACAGGAAAAATGATTAATCTCTAGTTGCTCGAGGACTT 5880  
Qy 5881 GGAGTCACTGCTGTGGACTTACTTTCACCCAACTGGTATGCTGATGGGGGTGGAGTT 5940  
Db 5881 GGAGTCACTGCTGTGGACTTACTTTCACCCAACTGGTATGCTGATGGGGGTGGAGTT 5940  
Qy 5941 CAAAGTACGACGAGGAAAAACATGTAAAGAAAGTAATCTCCCACTCAACCSGGGTACAT 6000  
Db 5941 CAAAGTACGACGAGGAAAAACATGTAAAGAAAGTAATCTCCCACTCAACCSGGGTACAT 6000  
Qy 6001 GGCACCTCTAGCCCTTACAAAGACTAGATCTCTCAAACTACATGAACCCCTCGGTACC 6060  
Db 6001 GGCACCTCTAGCCCTTACAAAGACTAGATCTCTCAAACTACATGAACCCCTCGGTACC 6060  
Qy 6061 CATATCGGCTGGTAAGCCTATTATACCAACCTCACTGGGCTCCATGAGTCTCGGCC 6120  
Db 6061 CATATCGGCTGGTAAGCCTATTATATACCAACCTCACTGGGCTCCATGAGTCTCGGCC 6120  
Qy 6121 CAAACCCCTACTAACTGTGGATATGCCCTCCCTGAACTTCARGCCATATGTTTCAATC 6180  
Db 6121 CAAACCCCTACTAACTGTGGATATGCCCTCCCTGAACTTCARGCCATATGTTTCAATC 6180  
Qy 6181 CTTGTACTGAAACAAATGGAAACAACTTCAGCACAGAAATAAACAACAACCTTCCTGTTAGTA 6240  
Db 6181 CTTGTACTGAAACAAATGGAAACAACTTCAGCACAGAAATAAACAACAACCTTCCTGTTAGTA 6240  
Qy 6241 GGACCTCTGTTTCCAAATSTGAAATAAACCCATACCTCAAACTCACTGTGTAAATTT 6300  
Db 6241 GGACCTCTGTTTCCAAATSTGAAATAAACCCATACCTCAAACTCACTGTGTAAATTT 6300  
Qy 6301 AGCAATACATACACACCACTCCCAATGCATCAGTGGTAACTCTCCCAACACAA 6360  
Db 6301 AGCAATACATACACACCACTCCCAATGCATCAGTGGTAACTCTCCCAACACAA 6360  
Qy 6361 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCCTATCGTGTGTTG 6420  
Db 6361 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCCTATCGTGTGTTG 6420  
Qy 6421 AATGGCTCTTCAGAAATCTATGCTTCTCTCATCTTAGTGCCGCCCYATGRCCATCTAC 6480  
Db 6421 AATGGCTCTTCAGAAATCTATGCTTCTCTCATCTTAGTGCCGCCCYATGRCCATCTAC 6480  
Qy 6481 ACTGACAGAGTTTATACAGTTATGTGCATATCTAGCCCGGCAACAAAGAGTACCATT 6540  
Db 6481 ACTGACAGAGTTTATACAGTTATGTGCATATCTAGCCCGGCAACAAAGAGTACCATT 6540  
Qy 6541 CTTCTCTTTTGTATAGGACGAGGAGTCTAGGTGCACTAGGTGCACTAGGTGCGGTATC 6600  
Db 6541 CTTCTCTTTTGTATAGGACGAGGAGTCTAGGTGCACTAGGTGCACTAGGTGCGGTATC 6600

Qy 6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660  
Db 6601 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660  
Qy 6661 GTGCGGCACTCCCTGGTCACTCTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTT 6720  
Db 6661 GTGCGGCACTCCCTGGTCACTCTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTT 6720  
Qy 6721 CAAATCGAAGAGCTTTAGACTGCTAACCCCTGABAGAGGGGGAACCTGTTTATTTT 6780  
Db 6721 CAAATCGAAGAGCTTTAGACTGCTAACCCCTGABAGAGGGGGAACCTGTTTATTTT 6780  
Qy 6781 GGGGAAGAACTCTGTATTATTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAA 6840  
Db 6781 GGGGAAGAACTCTGTATTATTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAA 6840  
Qy 6841 ATTCSAGATCGAATACAACTAKAGCAGGAGCTTCGAAACACTCTGGACCTCTGGGGCCTC 6900  
Db 6841 ATTCSAGATCGAATACAACTAKAGCAGGAGCTTCGAAACACTCTGGACCTCTGGGGCCTC 6900  
Qy 6901 CTCAGCCRAATGGATGCTGATCTCCCTTCTTAGAACCTCTAGCAGCTATAATATG 6960  
Db 6901 CTCAGCCRAATGGATGCTGATCTCCCTTCTTAGAACCTCTAGCAGCTATAATATG 6960  
Qy 6961 CTACTCTCTTTGGACCCCTGTATCTTTTACCTCTCTTGTAACTTCTCTCCAGATC 7020  
Db 6961 CTACTCTCTTTGGACCCCTGTATCTTTTACCTCTCTTGTAACTTCTCTCCAGATC 7020  
Qy 7021 GAAGCTGTAAATACAAATCGAGCCCAAGATGCACTCAAGACTCAAGACTTACCGCAGA 7080  
Db 7021 GAAGCTGTAAATACAAATCGAGCCCAAGATGCACTCAAGACTCAAGACTTACCGCAGA 7080  
Qy 7081 CCCCTGGACCCGCTGTAGCCCAAGATCTGATGTTAATGACATCAAAAGGCACCCCTCT 7140  
Db 7081 CCCCTGGACCCGCTGTAGCCCAAGATCTGATGTTAATGACATCAAAAGGCACCCCTCT 7140  
Qy 7141 GAGGAAATCTCAGCTGCAACCTCTACTACGCCCAATTCAGCAGGAGCAGTTAGAGC 7200  
Db 7141 GAGGAAATCTCAGCTGCAACCTCTACTACGCCCAATTCAGCAGGAGCAGTTAGAGC 7200  
Qy 7201 GGTSGTCGGCCAACTCCCAACAGCCTTTAGGTTTCTCTGTGAGATGGGGGACTGAGA 7260  
Db 7201 GGTSGTCGGCCAACTCCCAACAGCCTTTAGGTTTCTCTGTGAGATGGGGGACTGAGA 7260  
Qy 7261 GACAGGACTAGCTGGATTTCTAGGCTGATTAAGAAATCCYTAAGCCTAGSTGGGAAGGTG 7320  
Db 7261 GACAGGACTAGCTGGATTTCTAGGCTGATTAAGAAATCCYTAAGCCTAGSTGGGAAGGTG 7320  
Qy 7321 ACCACATCCACTTTAAACACGGGGCTTGCACTTAGTTCACCTGACCAATCAGAGAG 7380  
Db 7321 ACCACATCCACTTTAAACACGGGGCTTGCACTTAGTTCACCTGACCAATCAGAGAG 7380  
Qy 7381 CTCACATAAATGCTAATTTAGGCAAGACAGAGGTTAAAGAAATAGCCAATCATYATATG 7440  
Db 7381 CTCACATAAATGCTAATTTAGGCAAGACAGAGGTTAAAGAAATAGCCAATCATYATATG 7440  
Qy 7441 MTGAGACACACAGCAGGAGGGAACAATGATCGGGATATAAACCAAGTTCGAGCCGGCAA 7500  
Db 7441 MTGAGACACACAGCAGGAGGGAACAATGATCGGGATATAAACCAAGTTCGAGCCGGCAA 7500  
Qy 7501 CGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTTAC 7560  
Db 7501 CGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTTAC 7560  
Qy 7561 TCTATTAATCTTGCACTGCR 7582  
Db 7561 TCTATTAATCTTGCACTGCR 7582

RESULT 4  
AX027480  
LOCUS

AX027480 7582 bp DNA linear PAT 16-SRP-2000

DEFINITION Sequence 30 from Patent FR2788784.  
ACCESSION AX027480  
VERSION AX027480.1 GI:10188444  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1  
AUTHORS Mallet, F., Voisset, C. and Paranhos, B.G.  
JOURNAL Patent: FR 2788784-A 30 28-JUL-2000;  
BIO MERIEUX (FR)  
FEATURES  
source  
1..7582  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 100.0%; Score 7489.6; DB 2; Length 7582;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCC 60  
DB 1 CAACAATCGGGATATAAACCCAGGCATTCGAGCTGGCAACAGCAGCCCCCTTTGGGTCC 60  
QY 61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAATCTTGCNACTGCA 120  
DB 61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTCACTCTATTAATCTTGCNACTGCA 120  
QY 121 CTTCTTGCTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTACCGTCCACCACTGC 180  
DB 121 CTTCTTGCTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTACCGTCCACCACTGC 180  
QY 181 TGTTTGCCACCCAGCCANACCTCGCTGACTCCCATCCCTCTGGAATCTCGAGGGTGC 240  
DB 181 TGTTTGCCACCCAGCCANACCTCGCTGACTCCCATCCCTCTGGAATCTCGAGGGTGC 240  
QY 241 CGCTGTGCTCTGATCCAGCGCGCCCATTCGCGCTCCCAATTCGGCTTAAAGGCTTCG 300  
DB 241 CGCTGTGCTCTGATCCAGCGCGCCCATTCGCGCTCCCAATTCGGCTTAAAGGCTTCG 300  
QY 301 CATTTGNTCTGCA CGCTAAGTGCCTGGGTTTGTCTTAATTTGAGCTGAACACTTANTCACT 360  
DB 301 CATTTGNTCTGCA CGCTAAGTGCCTGGGTTTGTCTTAATTTGAGCTGAACACTTANTCACT 360  
QY 361 GGGTTCCATGGTCTCTCTGTGACCCACCGCTTCTAATAKAACTATAACACTTACCACA 420  
DB 361 GGGTTCCATGGTCTCTCTGTGACCCACCGCTTCTAATAKAACTATAACACTTACCACA 420  
QY 421 TGGCCCAAGATTCATTCTTGGAAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAAATAC 480  
DB 421 TGGCCCAAGATTCATTCTTGGAAATCCGTGAGGSCAACGAACTCCAGGTCAGAGAAATAC 480  
QY 481 GARGCTTGCCACATCTTGGAAAGCGGCTGCTACRGTCTTGAAGTGGTTTCCACCACATC 540  
DB 481 GARGCTTGCCACATCTTGGAAAGCGGCTGCTACRGTCTTGAAGTGGTTTCCACCACATC 540  
QY 541 TTTGGGAGCTCTGTGAGCAAGGACCCCGGTACATTTTGGGRCACCAAGGATCC 600  
DB 541 TTTGGGAGCTCTGTGAGCAAGGACCCCGGTACATTTTGGGRCACCAAGGATCC 600  
QY 601 MAAGTATGGGAAACCTTCCCGCAGACAAAACGCCCCCTTAAGCTATTCTGGAAAT 660  
DB 601 MAAGTATGGGAAACCTTCCCGCAGACAAAACGCCCCCTTAAGCTATTCTGGAAAT 660  
QY 661 TGGGAMCAATTTGACCTTCAGACACTAAGAAAGAAACGACTTATTTCTTCTGCACTGCC 720  
DB 661 TGGGAMCAATTTGACCTTCAGACACTAAGAAAGAAACGACTTATTTCTTCTGCACTGCC 720  
QY 721 GCCTGGCACTCTCTGAGGGGAAGTATAAATTAACCACTCTTACAGCTAGACYCTCTTTG 780

DB 721 GCCTGGCACTCTCTGAGGGGAAGTATAAATTAACCACTCTTACAGCTAGACYCTCTTTG 780  
QY 781 TAGAAAAGGCAAAATGGAGTGAAGTGCATGAAGTACAAAATCTTTCTTTTATTAGAGACAA 840  
DB 781 TAGAAAAGGCAAAATGGAGTGAAGTGCATGAAGTACAAAATCTTTCTTTTATTAGAGACAA 840  
QY 841 CTCACAATTATGTAAAGAGTGTGATTTATGCCCTCAGGAAGCTTTCAGAGTCTACCTCC 900  
DB 841 CTCACAATTATGTAAAGAGTGTGATTTATGCCCTCAGGAAGCTTTCAGAGTCTACCTCC 900  
QY 901 CTATCCAGCATCCCGAGCTCTTCCCAATTAAGGACCCCTTCAACCAATAGG 960  
DB 901 CTATCCAGCATCCCGAGCTCTTCCCAATTAAGGACCCCTTCAACCAATAGG 960  
QY 961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCATATTTCCCAAT 1020  
DB 961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCATATTTCCCAAT 1020  
QY 1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAAATTCGGCCAGCCAGAGTGCATGTCYT 1080  
DB 1021 TATGACCCCTCCCAAGCAGTGGGAGGAAGAAATTCGGCCAGCCAGAGTGCATGTCYT 1080  
QY 1081 TTTTYYTCTCCAGACTTAAGCAAAATAAAAACAGACTTAGGTAAATTTCTCAGATAAYCCT 1140  
DB 1081 TTTTYYTCTCCAGACTTAAGCAAAATAAAAACAGACTTAGGTAAATTTCTCAGATAAYCCT 1140  
QY 1141 GATGGCTATATTGRTCTTTTACAAGGGTTAGGACAATCTTTGATCTGACATCGAGAGAT 1200  
DB 1141 GATGGCTATATTGRTCTTTTACAAGGGTTAGGACAATCTTTGATCTGACATCGAGAGAT 1200  
QY 1201 ATATATGTCACTGCTAAATCAGACACTAACCCCAATGAGAGAAAGTCCACCAATAACTGC 1260  
DB 1201 ATATATGTCACTGCTAAATCAGACACTAACCCCAATGAGAGAAAGTCCACCAATAACTGC 1260  
QY 1261 AGCCTGAGGTTTGGCGATCTCTGATCTCAGTCAAGTCAATGGATANGATGACAAACA 1320  
DB 1261 AGCCTGAGGTTTGGCGATCTCTGATCTCAGTCAAGTCAATGGATANGATGACAAACA 1320  
QY 1321 GAAGGAAGCAANAATGATTTCCCAAGGCCAGCAGCAGTTCCAGCTCAGACCTCATTTG 1380  
DB 1321 GAAGGAAGCAANAATGATTTCCCAAGGCCAGCAGCAGTTCCAGCTCAGACCTCATTTG 1380  
QY 1381 GGGACACAGAAATCAGTAAACATGGGAGATTGGTGTCTGACAGACATTTGCTAACTTGTGTC 1440  
DB 1381 GGGACACAGAAATCAGTAAACATGGGAGATTGGTGTCTGACAGACATTTGCTAACTTGTGTC 1440  
QY 1441 TASAAGGACTAAGGAAAACCTASGAAGAAATCTAYGAATTACTCAATGATGCCACATA 1500  
DB 1441 TASAAGGACTAAGGAAAACCTASGAAGAAATCTAYGAATTACTCAATGATGCCACATA 1500  
QY 1501 ACACAGGGGAAGGGAAGAAATCCTACTGCTTCTGAGAGACTTAAGGAGGCAATTGAG 1560  
DB 1501 ACACAGGGGAAGGGAAGAAATCCTACTGCTTCTGAGAGACTTAAGGAGGCAATTGAG 1560  
QY 1561 GAAGGCTGCTCTCTGTCACTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTT 1620  
DB 1561 GAAGGCTGCTCTCTGTCACTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAGTT 1620  
QY 1621 TATCACTCAGTGCAGACATTTAGAAAAAATTTCAAAAGTCTGCCGTAGGCCCGGAG 1680  
DB 1621 TATCACTCAGTGCAGACATTTAGAAAAAATTTCAAAAGTCTGCCGTAGGCCCGGAG 1680  
QY 1681 CAAAACCTTAGAACCTATTGACCTTGGCACTTGGCACTCGGTTTTTATAATAGATCAGGAG 1740  
DB 1681 CAAAACCTTAGAACCTATTGACCTTGGCACTTGGCACTCGGTTTTTATAATAGATCAGGAG 1740  
QY 1741 GAGCAGGGGGAACAGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCAAGACCT 1800  
DB 1741 GAGCAGGGGGAACAGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCAAGACCT 1800  
QY 1801 CAGGCAAGTGGACTTTTGGAGGCTCTGGAAAAGGGAAAAGCTGGGCAAAATTTGAATGCTTAA 1860

Db 1801 CAGGCAAGTGGAGCTCTGGAGGCTCTGGAAGGGAAGAGCTGGGCAAAATGAAATGCCTAA 1860  
Qy 1861 TAGGCGTCTGCTTCCAGTGGGCTCTCAAGGACACTTTTAAAAAGATTTGTCCAGTAGAAG 1920  
Db 1861 TAGGCGTCTGCTTCCAGTGGGCTCTCAAGGACACTTTTAAAAAGATTTGTCCAGTAGAAG 1920  
Qy 1921 TAAGCGCGCCCTCTGCTCCATGCGCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCC 1980  
Db 1921 TAAGCGCGCCCTCTGCTCCATGCGCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCC 1980  
Qy 1981 CAGGGGCAAAAGGCTCTTTTGAAGTCAGAGGCCAATAACAGATGATCCAGCAGCAGGACTG 2040  
Db 1981 CAGGGGCAAAAGGCTCTTTTGAAGTCAGAGGCCAATAACAGATGATCCAGCAGCAGGACTG 2040  
Qy 2041 AGGGTCCCTGGGGCAAGCGCCATCCATGCGCATCCCTCACAGAGCCCTGGGTATGCTT 2100  
Db 2041 AGGGTCCCTGGGGCAAGCGCCATCCATGCGCATCCCTCACAGAGCCCTGGGTATGCTT 2100  
Qy 2101 GACCAATGAGGGCCAGGAAGGTTGTCTCTGACACCTGGTGGGCTTTCTTAGTCTTACT 2160  
Db 2101 GACCAATGAGGGCCAGGAAGGTTGTCTCTGACACCTGGTGGGCTTTCTTAGTCTTACT 2160  
Qy 2161 CTTCTGTCGCGACAACTGTCTCTCAGATCTGTCACTATTCTGAGGGGTCCTNTAAGACG 2220  
Db 2161 CTTCTGTCGCGACAACTGTCTCTCAGATCTGTCACTATTCTGAGGGGTCCTNTAAGACG 2220  
Qy 2221 GCGAGTCACTAGATCTTTTCCAGCCCACTAAGTTATGAAGTGGGAGCTTTTATCTTTT 2280  
Db 2221 GCGAGTCACTAGATCTTTTCCAGCCCACTAAGTTATGAAGTGGGAGCTTTTATCTTTT 2280  
Qy 2281 TCACATGCTTTTCTAATATGCTTGAAGCCCACTACCTTTGAGGAGAGACATTTCTA 2340  
Db 2281 TCACATGCTTTTCTAATATGCTTGAAGCCCACTACCTTTGAGGAGAGACATTTCTA 2340  
Qy 2341 GCAAGAGAGGGGCCAATTATACCTGAAATAGAGAGGAACACCCGTTGTTGTCNC 2400  
Db 2341 GCAAGAGAGGGGCCAATTATACCTGAAATAGAGAGGAACACCCGTTGTTGTCNC 2400  
Qy 2401 CTTGCTTGAGGAAGGAATTAATCTGAACTCTGGGCAACAGAAAGCAATATGGAAGCAGC 2460  
Db 2401 CTTGCTTGAGGAAGGAATTAATCTGAACTCTGGGCAACAGAAAGCAATATGGAAGCAGC 2460  
Qy 2461 CAAGAATGCCCCGCTCTGTTCAAGTTAACTTAAGGATTCACCTTCCCTACCAA 2520  
Db 2461 CAAGAATGCCCCGCTCTGTTCAAGTTAACTTAAGGATTCACCTTCCCTACCAA 2520  
Qy 2521 GGCAGTACCCCTCAGACCAAGGCCCAACAGGATTCCAAAGATTTGTTAAGGACTTAA 2580  
Db 2521 GGCAGTACCCCTCAGACCAAGGCCCAACAGGATTCCAAAGATTTGTTAAGGACTTAA 2580  
Qy 2581 AAGCCCAAGGCTTAGTAAACCATGATACTCCCTGCAAGTAATCCGTTAGGATTTGAG 2640  
Db 2581 AAGCCCAAGGCTTAGTAAACCATGATACTCCCTGCAAGTAATCCGTTAGGATTTGAG 2640  
Qy 2641 GAGGCAAGAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTTCAATGG 2700  
Db 2641 GAGGCAAGAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTTCAATGG 2700  
Qy 2701 AGGCGGTTGCTTTTATACCCAGCTGTACTAGCCCTTATACCTGTGMYTTCCCAAATAC 2760  
Db 2701 AGGCGGTTGCTTTTATACCCAGCTGTACTAGCCCTTATACCTGTGMYTTCCCAAATAC 2760  
Qy 2761 CAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAAGATGCCCTTCTGCACTCCCTG 2820  
Db 2761 CAGAGGAAGCAGAGTGGTTTACASTCCTGGACCTTMAAGATGCCCTTCTGCACTCCCTG 2820  
Qy 2821 TACATCTGACTCAATCTTTGTTGCTTTGAGATTTCAAAACCCARCATCTCAAC 2880  
Db 2821 TACATCTGACTCAATCTTTGTTGCTTTGAGATTTCAAAACCCARCATCTCAAC 2880  
Qy 2881 TCACCTGGAAGTCTTTTACCCCAAGGGTTTCAAGGATAGYCCCCATCTATTGTCAGGCAAT 2940  
Db 2881 TCACCTGGAAGTCTTTTACCCCAAGGGTTTCAAGGATAGYCCCCATCTATTGTCAGGCAAT 2940

Qy 2941 TAGCCCAAGACTTGAGYCAPTYMTACATCTGGACACTCTTGTCTTCTCRGTAKGTGGATG 3000  
Db 2941 TAGCCCAAGACTTGAGYCAPTYMTACATCTGGACACTCTTGTCTTCTCRGTAKGTGGATG 3000  
Qy 3001 ATTTACTTTTTRGCGCCYRTTCAGAAACCTTGTGCGCATCAAGCCACCCAAAGCRCTCTTAA 3060  
Db 3001 ATTTACTTTTTRGCGCCYRTTCAGAAACCTTGTGCGCATCAAGCCACCCAAAGCRCTCTTAA 3060  
Qy 3061 ATTTCTCGCYACCTGTGGCTTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGC 3120  
Db 3061 ATTTCTCGCYACCTGTGGCTTACAWGGTTTCCAAACSARARGCTCARCTCTGCTCACAGC 3120  
Qy 3121 AGGTTAAATACTTAGGCTAARATTTATCCAAAGGACCARGCCCTCAGTGAAGGAYRYA 3180  
Db 3121 AGGTTAAATACTTAGGCTAARATTTATCCAAAGGACCARGCCCTCAGTGAAGGAYRYA 3180  
Qy 3181 TCCAGCCTTACTGCTTATCTCATCYCAAAACCTTAAAGCAACTAAGRGRTTTCCTTG 3240  
Db 3181 TCCAGCCTTACTGCTTATCTCATCYCAAAACCTTAAAGCAACTAAGRGRTTTCCTTG 3240  
Qy 3241 GCRTAAAYAGGYTTCTGCCGAATWATGGATTCCCGAGTWTGCGRAAATAGCCAGGYCATTA 3300  
Db 3241 GCRTAAAYAGGYTTCTGCCGAATWATGGATTCCCGAGTWTGCGRAAATAGCCAGGYCATTA 3300  
Qy 3301 WATACASTAATTAAGGAACCTCAGAAAGCCCAATACCCATTTTARTAAGATGGAYMCTGAA 3360  
Db 3301 WATACASTAATTAAGGAACCTCAGAAAGCCCAATACCCATTTTARTAAGATGGAYMCTGAA 3360  
Qy 3361 GYMAAGTGGCTTTCAGGCCCTTAAAGAGGCCCTTAAACCCAGYCCAGTGTTAAGYT 3420  
Db 3361 GYMAAGTGGCTTTCAGGCCCTTAAAGAGGCCCTTAAACCCAGYCCAGTGTTAAGYT 3420  
Qy 3421 TGCCAAACRGGCAAGACTTTTSTYATAYRTCAAGAAAAAACAAGRAAYAGCTCTRGA 3480  
Db 3421 TGCCAAACRGGCAAGACTTTTSTYATAYRTCAAGAAAAAACAAGRAAYAGCTCTRGA 3480  
Qy 3481 GTCTTTACACAGRCCBRAGGAYGAGCTTCAACCCYRTGGCRYACCTCASTAAAGAAAT 3540  
Db 3481 GTCTTTACACAGRCCBRAGGAYGAGCTTCAACCCYRTGGCRYACCTCASTAAAGAAAT 3540  
Qy 3541 GATGTAGTGGCAAGGGTGTGCTTATTTTAYGGGTAGTGGTGGCAGTACAGTYKTA 3600  
Db 3541 GATGTAGTGGCAAGGGTGTGCTTATTTTAYGGGTAGTGGTGGCAGTACAGTYKTA 3600  
Qy 3601 GTATCTGAAGCAGTTTAAATAATAACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
Db 3601 GTATCTGAAGCAGTTTAAATAATAACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
Qy 3661 GTGAAYRGCATCTACTCCTGCTTAAAGGAGACTTGTGGCTGTGACAAACYGTTTACTTAA 3720  
Db 3661 GTGAAYRGCATCTACTCCTGCTTAAAGGAGACTTGTGGCTGTGACAAACYGTTTACTTAA 3720  
Qy 3721 TRTCAAGGCTCTATTACTTGAAGGCCAGTGTGCTGCRACCTGTGCACTTGTGCAACTCTTAAC 3780  
Db 3721 TRTCAAGGCTCTATTACTTGAAGGCCAGTGTGCTGCRACCTGTGCACTTGTGCAACTCTTAAC 3780  
Qy 3781 CCAGYCNCAITTTCTTCCAGACAAATGAAGAAAGATARAAYATAAAGTGTGCAACAAATTAAT 3840  
Db 3781 CCAGYCNCAITTTCTTCCAGACAAATGAAGAAAGATARAAYATAAAGTGTGCAACAAATTAAT 3840  
Qy 3841 TCTCAAACTATGCCACTCGAGGGGACCTTTTAGARGTTCCYTTGACATCGYACCTT 3900  
Db 3841 TCTCAAACTATGCCACTCGAGGGGACCTTTTAGARGTTCCYTTGACATCGYACCTT 3900  
Qy 3901 CAACTGTGTATCTGATGGAAGTCTCTTGTAGAAAAAGGACTTCGAAAAAGYGGGTATGC 3960  
Db 3901 CAACTGTGTATCTGATGGAAGTCTCTTGTAGAAAAAGGACTTCGAAAAAGYGGGTATGC 3960  
Qy 3961 AGTGTGCTAGTGAATAATGGAATTTGAAAGTAATCCCTCACTCCAGGAACCTAGTCTTAA 4020  
Db 3961 AGTGTGCTAGTGAATAATGGAATTTGAAAGTAATCCCTCACTCCAGGAACCTAGTCTTAA 4020

QY 4021 GCTRGCAACTAATAGCCYTCAYTKGGCACTAGAAATAGGAGAAAGGAAAGGAA 4080  
DB 4021 GCTRGCAACTAATAGCCYTCAYTKGGCACTAGAAATAGGAGAAAGGAAAGGAA 4080  
QY 4081 TATATATACAGACTCTRTARTATGCTYACCTAGTCNTCCATGMRGCAATATGSA 4140  
DB 4081 TATATATACAGACTCTRTARTATGCTYACCTAGTCNTCCATGMRGCAATATGSA 4140  
QY 4141 AGAAGGGAATTCCTAATCTCYGAGGAAACACCTATCATCAGAAAGCCATTAGGAR 4200  
DB 4141 AGAAGGGAATTCCTAATCTCYGAGGAAACACCTATCATCAGAAAGCCATTAGGAR 4200  
QY 4201 ATTATTAATGCGGTACAGAAACCTARAGAGGTGGMAGTCTTAACCTGCGGGTCA 4260  
DB 4201 ATTATTAATGCGGTACAGAAACCTARAGAGGTGGMAGTCTTAACCTGCGGGTCA 4260  
QY 4261 NAAAGGAAAGAAAGGAAATASAAGRAAYTGCACAGCAKATATTGAAGCMAAAGAGC 4320  
DB 4261 NAAAGGAAAGAAAGGAAATASAAGRAAYTGCACAGCAKATATTGAAGCMAAAGAGC 4320  
QY 4321 TGCAAGGCAAGCCCTCCATTAGAAATGCTTATTAACCTTCCCTTAGTATAGGTTATCC 4380  
DB 4321 TGCAAGGCAAGCCCTCCATTAGAAATGCTTATTAACCTTCCCTTAGTATAGGTTATCC 4380  
QY 4381 CTTCCGGGAAACCAAGCCCACTACTCAGCAGAGAAACAGAAATGSGGAACCTCAGAGG 4440  
DB 4381 CTTCCGGGAAACCAAGCCCACTACTCAGCAGAGAAACAGAAATGSGGAACCTCAGAGG 4440  
QY 4441 CAGTTTTCTCCCTCGGACGGTTAGCCACTGAAGAGGAAATACTTTTGCTGCAAC 4500  
DB 4441 CAGTTTTCTCCCTCGGACGGTTAGCCACTGAAGAGGAAATACTTTTGCTGCAAC 4500  
QY 4501 TATCCAAATGAAATTAATCTTAAACCCCTTCAATCAAACTTTCACTTAGGCAATAGCAC 4560  
DB 4501 TATCCAAATGAAATTAATCTTAAACCCCTTCAATCAAACTTTCACTTAGGCAATAGCAC 4560  
QY 4561 CCATCARATGSCCAATCATTTATTAAGTACAGAGGCTTTTCAAACTATCAAGCARAT 4620  
DB 4561 CCATCARATGSCCAATCATTTATTAAGTACAGAGGCTTTTCAAACTATCAAGCARAT 4620  
QY 4621 AKTCAGGCGCTGTGAATGTGCCARARAATAATCCCTGCTATGCGCAAGCTCTTC 4680  
DB 4621 AKTCAGGCGCTGTGAATGTGCCARARAATAATCCCTGCTATGCGCAAGCTCTTC 4680  
QY 4681 AGAARAAARAAACAGGCAATACCTGGRARAARACTGGCAACTGATTTTACCCACAAG 4740  
DB 4681 AGAARAAARAAACAGGCAATACCTGGRARAARACTGGCAACTGATTTTACCCACAAG 4740  
QY 4741 CCCAAACCTCAGGATTTCACTAGTCTGGTATATATCTTTCAGGGTTGGGCA 4800  
DB 4741 CCCAAACCTCAGGATTTCACTAGTCTGGTATATATCTTTCAGGGTTGGGCA 4800  
QY 4801 RAGGCTTCCCTGTAGGACAGAAAGGCCCCAAGAGTAAAGGCACTAGTTTCATGAA 4860  
DB 4801 RAGGCTTCCCTGTAGGACAGAAAGGCCCCAAGAGTAAAGGCACTAGTTTCATGAA 4860  
QY 4861 ATAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTCCAG 4920  
DB 4861 ATAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTCCAG 4920  
QY 4921 GCCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATAGATATCACTTACACTCGGC 4980  
DB 4921 GCCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATAGATATCACTTACACTCGGC 4980  
QY 4981 TGAAGCCACAGTCCTCAGGGAAGGTCGAAATGAATGAAYACTCAAGGACATCTA 5040  
DB 4981 TGAAGCCACAGTCCTCAGGGAAGGTCGAAATGAATGAAYACTCAAGGACATCTA 5040  
QY 5041 AAAAAGCAACCCAGGAAACCCACCTCACAATGCGCTGCTGTTGCTATAGCCTTAA 5100  
DB 5041 AAAAAGCAACCCAGGAAACCCACCTCACAATGCGCTGCTGTTGCTATAGCCTTAA 5100  
QY 5101 AGAATCTGCAACTTTTCCCAAAAGCAGGACTTAGGCCATACGAAATGCTGTATGGAAG 5160

DB 5101 AGAATCTGCAACTTTTCCCAAAAGCAGGACTTAGGCCATACGAAATGCTGTATGGAAG 5160  
QY 5161 CCCTTCATACCAATGACCTTGTGCTTGACCCAAAGCAGCAAACTTAGTTGCGACATCA 5220  
DB 5161 CCCTTCATACCAATGACCTTGTGCTTGACCCAAAGCAGCAAACTTAGTTGCGACATCA 5220  
QY 5221 CCTCTCTAGCCAAATATCAACAAAGTCTTAAACATTTACAAAGAACCTATCCTCTAGA 5280  
DB 5221 CCTCTCTAGCCAAATATCAACAAAGTCTTAAACATTTACAAAGAACCTATCCTCTAGA 5280  
QY 5281 AGGGAAGAAAGTATTTCCACCCWGTGACATGTTAGTCAAGTCCCTCYCTCTAAT 5340  
DB 5281 AGGGAAGAAAGTATTTCCACCCWGTGACATGTTAGTCAAGTCCCTCYCTCTAAT 5340  
QY 5341 CCCCATCTCTAGATACATCTCTGGGAAAGGACCTTACCAGCTCATTTTATYACCCCACTG 5400  
DB 5341 CCCCATCTCTAGATACATCTCTGGGAAAGGACCTTACCAGCTCATTTTATYACCCCACTG 5400  
QY 5401 CGTTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAATCTTGGATAC 5460  
DB 5401 CGTTTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAATCTTGGATAC 5460  
QY 5461 TGCCAAAGGAACCTGAAATCCAGGAGACAACTAGCTATTCTCTGTAACCTCTAGAGG 5520  
DB 5461 TGCCAAAGGAACCTGAAATCCAGGAGACAACTAGCTATTCTCTGTAACCTCTAGAGG 5520  
QY 5521 ATTTGCGCTGCTCTTCAAAACCAACAGGAGGAAAGTAACTAAAATCATAAATCCCCC 5580  
DB 5521 ATTTGCGCTGCTCTTCAAAACCAACAGGAGGAAAGTAACTAAAATCATAAATCCCCC 5580  
QY 5581 ATGSGCCTCCTTATCATATTTTCTKTASTGTTSTTTTACCCTSTTTCACTCTCACT 5640  
DB 5581 ATGSGCCTCCTTATCATATTTTCTKTASTGTTSTTTTACCCTSTTTCACTCTCACT 5640  
QY 5641 GCACCCCTCCATGCGCTGTATGACAGTAGCTCCCTYACCMAGAGTTTCTATGAGA 5700  
DB 5641 GCACCCCTCCATGCGCTGTATGACAGTAGCTCCCTYACCMAGAGTTTCTATGAGA 5700  
QY 5701 ATGACAGCTCCCGAAATATTGATGCCCATCTATAGAGTCTTTTSTAAGGGAACCC 5760  
DB 5701 ATGACAGCTCCCGAAATATTGATGCCCATCTATAGAGTCTTTTSTAAGGGAACCC 5760  
QY 5761 ACCTTCACTGCCACACCCATATATGCGCCGCACTGCTATCACTCTGCGACTCTTTCATG 5820  
DB 5761 ACCTTCACTGCCACACCCATATATGCGCCGCACTGCTATCACTCTGCGACTCTTTCATG 5820  
QY 5821 CATGCAAAATCACTATTATTGGACAGGAAATAATTAATCTCTAGTTGCTCTGAGGACTT 5880  
DB 5821 CATGCAAAATCACTATTATTGGACAGGAAATAATTAATCTCTAGTTGCTCTGAGGACTT 5880  
QY 5881 GGAGTCACTGCTGTTGGACTTACTTACCCAACTGGTATGCTCATGCGGGTGGAGTT 5940  
DB 5881 GGAGTCACTGCTGTTGGACTTACTTACCCAACTGGTATGCTCATGCGGGTGGAGTT 5940  
QY 5941 CAAGATCAGGCAAGGAAACATGTAAAGAAAGTAACTCTCCAACTCACCCSGGTACAT 6000  
DB 5941 CAAGATCAGGCAAGGAAACATGTAAAGAAAGTAACTCTCCAACTCACCCSGGTACAT 6000  
QY 6001 GGCACTCTTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAAACCTCTCCGTACC 6060  
DB 6001 GGCACTCTTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAAACCTCTCCGTACC 6060  
QY 6061 CATACTCGCTGGTAAAGCTTAAATACCCCTCAGCTGCGGTCCATGAGTCTCGGC 6120  
DB 6061 CATACTCGCTGGTAAAGCTTAAATACCCCTCAGCTGCGGTCCATGAGTCTCGGC 6120  
QY 6121 CAAAACCTTACTTAACTGTTGGATATGCTCTCCCTGAACTTCARGGCAATATGTTTCAATC 6180  
DB 6121 CAAAACCTTACTTAACTGTTGGATATGCTCTCCCTGAACTTCARGGCAATATGTTTCAATC 6180  
QY 6181 CTTGTAACCTGAA CAATGGAAACAACTTACGACAGAAATAAAACCACTTCCGTTTTAGTA 6240



Db 6181 CCTGTACTGTAACAATGGAACAACCTTCAGCACAGAAAATAAACACACACTTCCTCGTTTGTAGTA 6240  
Qy 6241 GGACCTCTGTTTCCAAATSTGAATAAACCCATACCTCAAACTCACCTGTGTAAATTT 6300  
Db 6241 GGACCTCTGTTTCCAAATSTGAATAAACCCATACCTCAAACTCACCTGTGTAAATTT 6300  
Qy 6301 AGCAATATACATACACAACCAACTCCCAATGATCAGGTGGTAACTCTCCACACAA 6360  
Db 6301 AGCAATATACATACACAACCAACTCCCAATGATCAGGTGGTAACTCTCCACACAA 6360  
Qy 6361 ATAGTCTGCTACCTTACAGGAATATTTTGTCTGTGTACCTCAGCCCTATCGTTGTTG 6420  
Db 6361 ATAGTCTGCTACCTTACAGGAATATTTTGTCTGTGTACCTCAGCCCTATCGTTGTTG 6420  
Qy 6421 AATGCTCTTCAAGATCTATGCTTCCTCTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6480  
Db 6421 AATGCTCTTCAAGATCTATGCTTCCTCTCATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6480  
Qy 6481 ACTGAACAAGATTATACAGTTATGTCTATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540  
Db 6481 ACTGAACAAGATTATACAGTTATGTCTATATCTAAGCCCCGCAACAAAAGAGTACCCATT 6540  
Qy 6541 CTTCTTTTGTATAGGACGAGGTGCTAGTGCACTAGTACTGGCAATGGCGGTATC 6600  
Db 6541 CTTCTTTTGTATAGGACGAGGTGCTAGTGCACTAGTACTGGCAATGGCGGTATC 6600  
Qy 6601 ACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660  
Db 6601 ACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAACGG 6660  
Qy 6661 GTCCGGCAGTCCCTGCTGCTCAAGATCAACTTAACTCCCTAGCAGCAGTGCCTT 6720  
Db 6661 GTCCGGCAGTCCCTGCTGCTCAAGATCAACTTAACTCCCTAGCAGCAGTGCCTT 6720  
Qy 6721 CRAAATCGAAGAGCTTTAGACTGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTA 6780  
Db 6721 CRAAATCGAAGAGCTTTAGACTGCTAACCGCTGARAGAGGGGGAACCTGTTTATTTTA 6780  
Qy 6781 GGGGAAGATGCTGTTATATGTTTAACTCAATCGGAATCGTCACTGAGAAAGTTAAAGAA 6840  
Db 6781 GGGGAAGATGCTGTTATATGTTTAACTCAATCGGAATCGTCACTGAGAAAGTTAAAGAA 6840  
Qy 6841 ATTCSAGATCGAATCAACAGTAKAGCAGAGAGCTTCGAAACACTGGACCTGGGGCCTC 6900  
Db 6841 ATTCSAGATCGAATCAACAGTAKAGCAGAGAGCTTCGAAACACTGGACCTGGGGCCTC 6900  
Qy 6901 CTCAGCCRATGGATGCCCTGGATTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATG 6960  
Db 6901 CTCAGCCRATGGATGCCCTGGATTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATG 6960  
Qy 6961 CTACTCCTCTTTGGACCTGATCTTTTRACCTCCTTGTAACTTTGTCTCTCCAGAAATC 7020  
Db 6961 CTACTCCTCTTTGGACCTGATCTTTTRACCTCCTTGTAACTTTGTCTCTCCAGAAATC 7020  
Qy 7021 GAAGCTGTAAACTCAAAATGAGCCCAAGATGCAAGTCAAGACTAAAGACTACCGCAGA 7080  
Db 7021 GAAGCTGTAAACTCAAAATGAGCCCAAGATGCAAGTCAAGACTAAAGACTACCGCAGA 7080  
Qy 7081 CCCCTGGACCGGCTGYTAGCCACGATCTGATGTTAATGATCAAAAGGACCCCTCCT 7140  
Db 7081 CCCCTGGACCGGCTGYTAGCCACGATCTGATGTTAATGATCAAAAGGACCCCTCCT 7140  
Qy 7141 GAGGAAATCTAGCTGCAACACTCTATACGCCCAATTCAGCAGGAGCAGTTAGAGC 7200  
Db 7141 GAGGAAATCTAGCTGCAACACTCTATACGCCCAATTCAGCAGGAGCAGTTAGAGC 7200  
Qy 7201 GGTSGTGGCCCAACTCCCAACAGCAGTCTAGCTTTCTGTTGATGAGTGGGGACTGAGA 7260  
Db 7201 GGTSGTGGCCCAACTCCCAACAGCAGTCTAGCTTTCTGTTGATGAGTGGGGACTGAGA 7260  
Qy 7261 GACGAGCTAGCTGATTTCTTAGGCTGATTAAGAAATCCYTAAGCTAGTGGGAAGTGT 7320  
Db 7261 GACGAGCTAGCTGATTTCTTAGGCTGATTAAGAAATCCYTAAGCTAGTGGGAAGTGT 7320

Qy 7321 ACCCATCCACTTTAAACACGGGCTTGCACCTTAGTTCACACCTGACCAATCAGAGAG 7380  
Db 7321 ACCCATCCACTTTAAACACGGGCTTGCACCTTAGTTCACACCTGACCAATCAGAGAG 7380  
Qy 7381 CTCCTAAAATGCTAAATTTAGGCAAAAGACAGGAGGTAAAGAAATAGCCAAATCATYATTGC 7440  
Db 7381 CTCCTAAAATGCTAAATTTAGGCAAAAGACAGGAGGTAAAGAAATAGCCAAATCATYATTGC 7440  
Qy 7441 MTGAGAGCACAGCAGGAGGGAACAATGATCGGATATAAACCAGTTCGAGCCGGCAA 7500  
Db 7441 MTGAGAGCACAGCAGGAGGGAACAATGATCGGATATAAACCAGTTCGAGCCGGCAA 7500  
Qy 7501 CGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCAC 7560  
Db 7501 CGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCAC 7560  
Qy 7561 TCTATTAATCTTGCARCTGCR 7582  
Db 7561 TCTATTAATCTTGCARCTGCR 7582  
RESULT 5  
LOCUS AY101582  
DEFINITION Homo sapiens isolate 21 endogenous retrovirus HERV-W, ERVWE1 locus, allele A, complete sequence.  
ACCESSION AY101582  
VERSION AY101582.1 GI:37544399  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 10222)  
AUTHORS Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.  
TITLE The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)  
PUBMED 14757826  
REFERENCE 2 (bases 1 to 10222)  
AUTHORS Mallet, F., Bouton, O. and Oriol, G.  
TITLE Direct Submission  
JOURNAL Submitted (06-MAY-2002) Retrovirology Department, UMR 2142 CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France  
FEATURES  
source  
1..10222  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/isolate="21"  
/isolation\_source="PBMC of individual 21"  
/db\_xref="taxon:9606"  
/chromosome="7"  
/map="7q21-q22"  
/sex="female"  
/cell\_type="PBMC"  
/note="isolated by PCR; ERVWE1 locus, allele A endogenous\_virus: HERV-W"  
1..780  
join(248..811,2887..9813)  
/note="putative mRNA transcript 1; gag-pol pseudogene"  
join(248..811,7585..9813)  
/product="envelope glycoprotein"  
/note="putative mRNA transcript 2"  
join(248..811,9248..9813)  
/note="putative mRNA transcript 3; no coding region"  
7812..9428  
/note="syncytin"  
/codon\_start=1  
/product="envelope glycoprotein"



```

/protein_id="AM68161.1"
/db_xref="GI:3754440"
/translation="MALPVIHFLFTVLLPSFTLAPPCCWTSSTSSPQSFLEWMQRP
GNIDAPYSRLSKGTPTFAHTMERNYCHSATCLMHANTHWYTKMINSCGGLGV
TWCWYPTQTGMSDGGVQDQAREKHVEVLSQPLRVHGSTSPYKGLDLSKLHETRT
HTRLSVLSVDSVQQAQNTWCICLPLNFRPYSVSPVQNMNFETINTTSV
LVGFLPVNLSIHTSNLCATKSFNTYTNQCIKRVPTQTIVCLSPSPFVCGTSA
YRCLNGSSSMCFPLSPVLPMTIYTEQLSVKTSKPNKRPVTLPIVIGAVLGALG
TGIGLITSTQFYKLSGLNGDMERVDASLVTLQDOLNSIAAVLVQNRALLDTAE
RGCTGLFLGEECCYYNQSGIVTEKVEIRDIRIQRAEELNQPWGLSKQMPWLP
FGFLAAIILLFLFGYPLFNLLIYVNSRIEAVKLQMEPQSKTKTYRPLDRPASP
RSDVNDIKGTPPEPTEISAAQFLRLPNSAGS"
9487..10222

```

Query Match	86.1%;	Score 6448;	DB 5;	Length 10222;
Best Local Similarity	95.7%;	Pred. No. 0;		
Matches 6681;	Conservative 195;	Mismatches 50;	Indels 53;	Gaps 27;
QY	606	GATGGGAAACGTTCCCGCAAGACAAAAACGCCCCCTAAGACGTAATCTGCGARAAATTCGGGA	665	
Db	2886	GATGGGAAACGTTCCCGCAAGACAAAAACGCCCCCTAAGACGTAATCTGCGARAAATTCGGGA	2945	
QY	666	MCAATTGTGACCCCTCAGACACTAAAGAAAGAAACGACTTATATCTTCTGCGAGTGCCTGCTG	725	
Db	2946	CCAATTGTGACCCCTCAGACACTAAAGAAAGAAACGACTTATATCTTCTGCGAGTGCCTGCTG	3005	
QY	726	GCACCTCTGAGGGAAGTATATAATTAACAACCACTTTTAACAGCTAGACATCTTTTGTAGAA	785	
Db	3006	GCACCTCTGAGGGAAGTATATAATTAACAACCACTTTTAACAGCTAGACATCTTTTGTAGAA	3065	
QY	786	AAGGCAAAATGGAGTGAAGTCCCAATAGTACAAACTTCTTTTCACTTAAGAGACAACCTCAC	845	
Db	3066	AAGGCAAAATGGAGTGAAGTCCCAATAGTACAAACTTCTTTTCACTTAAGAGACAACCTCAC	3125	
QY	846	AATTATGTAAAAAGTGTGATTTATATGCCCTACAGGAAGCCCTTCAGAGTCTACCTCCCTATC	905	
Db	3126	AATTATGTAAAAAGTGTGATTTATATGCCCTACAGGAAGCCCTTCAGAGTCTACCTCCCTATC	3185	
QY	906	CCAGCATCCCGGACTCCTTCCCCCAATTAATAAGGACCCGCCCTTCAACCCAAATGGTCCAA	965	
Db	3186	CCAGCATCCCGGACTCCTTCCCCCAATTAATAAGGACCCGCCCTTCAACCCAAATGGTCCAA	3245	
QY	966	AAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCTCAATATTCGCCCAATATATGA	1025	
Db	3246	AAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCTCAATATTCGCCCAATATATGA	3305	
QY	1026	CCCTCCCAAGCAGTGGGAGGAAGATAATTCGGCCCAAGCAGAGTGCATGTGCTTTTYY	1085	
Db	3306	CCCTCCCAAGCAGTGGGAGGAAGATAATTCGGCCCAAGCAGAGTGCATGTGCTTTTTC	3364	
QY	1086	TCTCCAGACCTTAAGCAAAATTAACAGACACTTAGGTAAATCTCAGATTAACCTGATGG	1145	
Db	3365	TCTCCAGACCTTAAGCAAAATTAACAGACACTTAGGTAAATCTCAGATTAACCTGATGG	3424	
QY	1146	CTATATTGRTGTTTTACAAGGGTTAGGACAAATCTTTTGATCTGACATGGAGAGATATATA	1205	
Db	3425	CTATATTGRTGTTTTACAAGGGTTAGGACAAATCTTTTGATCTGACATGGAGAGATATATA	3483	
QY	1206	TGTCACCTGCTAAATCAGACACTAACCCCAATCAGAGAAGTGCACCAATAACTGCAGCCT	1265	
Db	3484	TGTCACCTGCTAAATCAGACACTAACCCCAATCAGAGAAGTGCACCAATAACTGCAGCCT	3543	
QY	1266	GAGGTTTGGCGATCTCTGATATCTCAGTCAGGTCAATGGATANGATGACACAGAAG	1325	
Db	3544	GAGGTTTGGCGATCTCTGATATCTCAGTCAGGTCAATGGATANGATGACACAGAAG	3600	
QY	1326	AAAGANAATGATTTCCCAACAGGCCAGCAGCAGTTCCTCAGTCTASACCTTCATTTGGGGAC	1385	
Db	3601	AAAGANAATGATTTCCCAACAGGCCAGCAGCAGTTCCTCAGTCTASACCTTCATTTGGGGAC	3657	
QY	1386	ACAGAAATCAGTAACATGGGAGATTTGGTGTCTCAGACAATTTGCTAACTTTGTGTCTASAA	1445	

3658	DB	ACAGAAATCAGAACAT - GGAGATTGGTCTGCAGACATTTGCTAACTTTGTGTGCTAGAA	3711
1446	QY	GGACTAAGGAAACATASGAAGAAARTCTAYGAATTTACTCAATGATGTCACACCATAAACA	1505
3717	DB	GGACTAAGGAAACATAGGAG - AAGTCTAGTAATTTACTCAATGATGTCACACCATAAACA	3775
1506	QY	GGGGAAGGGAAGAAATCTCTACTCGCTTTCTGGAGAGACTAAAGGAGGCAATCAGGAAGC	1565
3776	DB	- GGGGAAGGGAAGAAATCTCTACTCGCTTTCTGGAGAGACTAAAGGAGGCAATCAGGAAGC	3834
1566	QY	GTGCCTCTCTGTCACCTGACTCTTCTGAGGCCAACTTAATCTTAAAGCGTTAGTTTATCA	1625
3835	DB	GTGCCTCTCTGTCACCTGACTCTTCTGAGGCCAACTTAATCTTAAAGCGTTAGTTTATCA	3894
1626	QY	CTCAGTCAGCTGCAGACATTAG - AAAAATCTTCAAAAGTCTGCGTAGGCCCCGGAGCAAA	1684
3895	DB	CTCAGTCAGCTGCAGACATTAGAAATAAAATCTTCAAAAGTCTGCGTAGGCCCCGGAGCAAA	3954
1685	QY	ACTTAGAAACCTTATTGAACCTTGGCAACTTCGGTTTTTATATAGAGATCAGGAGGAGC	1744
3955	DB	ACTTAGAAACCTTATTGAACCTTGGCAACCTTCGGTTTTTATATAGAGATCAGGAGGAGC	4014
1745	QY	AGCGGAAACAGGACAAACGGGATTTAAATAAAAGGCCACCGCTTTAGTCATGACCCCTCAGG	1804
4015	DB	AGCGGAAACAGGACAAACGGGATTTAAATAAAAGGCCACCGCTTTAGTCATGACCCCTCAGG	4074
1805	QY	CAAGTGACATTTGAGGCTCTGAAAAGGGAAGCTGGGCAATTTGAATGCCATAATAGG	1864
4075	DB	CAAGTGACATTTGAGGCTCTGAAAAGGGAAGCTGGGCAATTTGAATGCCATAATAGG	4134
1865	QY	GCTTGTCTCAGTGCAGTCTPAAGGACACTTTTAAAAAGATTTGTCCAAAGTAGAAGTAAG	1924
4135	DB	GCTTGTCTCAGTGCAGTCTPAAGGACACTTTTAAAAAGATTTGTCCAAAGTAGAAGTAAG	4194
1925	QY	CCGCCCTTCGTCATGCCCTTATTTCAAGGGAATCACTGGNAGGCCCACTGCCCCAGG	1984
4195	DB	CCGCCCTTCGTCATGCCCTTATTTCAAGGGAATCACTGGNAGGCCCACTGCCCCAGG	4254
1985	QY	GGACAAAGGTCTTTTGAGTCAGAGGCACTAAACAGATGATCCAGCAGCAGGACTGAGGG	2044
4255	DB	GGACAAAGGTCTCTGAGTCAGAGGCACTAAACAGATGATCCAGCAGCAGGACTGAGGG	4314
2045	QY	TGCCTGGGGCAAGCGCCATCCCATGCCATCACCTCACAGAGCCCTGGGTATGCTTGACC	2104
4315	DB	TGCCTGGGGCAAGCGCCATCCCATGCCATCACCTCACAGAGCCCTGGGTATGCTTGACC	4374
2105	QY	ATTGAGGGCCAGGAAGTTGTCTCTGGACACTGTGTGGCTCTTCTTAGTCTTACTCTTC	2164
4375	DB	ATTGAGGGCCAGG - AGGTTGTCTCTGGACACTGTGTGGCTCTTCTTAGTCTTACTCTTC	4433
2165	QY	TGTCCTGGGCAACTGTCTCCAGATCTGTCACTATTCTCAGGGGGTCCNTAAGACGGCA	2224
4434	DB	TGTCCTGGGCAACTGTCTCCAGATCTGTCACTA - TCTGAGGGGGTCC - TAAAGCGGGCA	4491
2225	QY	GTCACTAGATACTTCTTCCAGCCACTAAGTTATGAATCGGGAGCTTTTATCTTTTTCAC	2284
4492	DB	GTCACTAGATAC - TTTCTCCAGCCACTAAGTTATG - ACTGGGGAGCTTTTATCTTTTCAC	4549
2285	QY	ATGCTTTTCTAATATTGCTTGAAGGCCCACTACCTTGTATTAGGAGAGACATTCTAGCAA	2344
4550	DB	ATGCTTTTCTAATATTGCTTGAAGGCCCACTACCTTGTATTAGGAGAGACATTCTAGCAA	4609
2345	QY	AACGAGGGCCATTATACACCTGAAACATAGGAGAGAACACCGTTGTGTGTCNCCCTG	2404
4610	DB	AACGAGGGCCATTATACACCTGAAACATAGGAGAGAGAACACCGTTGTGTGTCNCCCTG	4668
2405	QY	CTTGAGGAAGGAATTAATCCTGAAAGTCTGGGCAACAGAGGACAAATATGGAGCGACAAA	2464
4669	DB	CTTGAGGAAGGAATTAATCCTGAAAGTCTGGGCAACAGAGGACAAATATGGAGCG - CAAA	4727
2465	QY	GAATGCCGCTCTGTTCAAGTTAAACTAAAGGATTCACCTTCTTCCCTACCAAGGCA	2524
4728	DB	GAATGCCGCTCTGTTCAAGTTAAACTAAAGGATTCACCTTCTTCCCTACCAAGGCA	4787

QY 2525 GTACCCCTCAGACCCCAAGGCCCAACAGGATTTCCAAAGATTGTTAAAGGACTTAAAGGC 2584  
DB GTACCCCTCAGACCCCAAGGCCCAACAGGATTTCCAAAGATTGTTAAAGGACTTAAAGGC 4847  
QY 2585 CCAAGGCTTAGTAAACCAATGATCAATCCCTCGAGTAAATCCGTAGTGGATTTAGAGGAG 2644  
DB CCAAGGCTTAGTAAACCAATGATCAATCCCTCGAGTAAATCCGTAGTGGATTTAGAGGAG 4900  
QY 2645 CACAGAAACCAAGTGGAGTGGAGGGTTAGTGGAGATCTCAGGATTTATCAATGGAGGC 2704  
DB TACAGAAACCAACAGACAGTGGGA--GGTTAGTGGAGATCTCAGGATTTATCAAT--GAGGC 4958  
QY 2705 GGTGCTCTTTATACCCAGCTGATACCTAGCCCTTATCTGTMVTTCCCAATACCAGA 2764  
DB TGTGTTCTCTATAGCAGCTGTACCTAGCCCTTATCTGCTTCCCAATACCAGA 5018  
QY 2765 GGAACAGAGTGGTTTACASTCTCGACCTTMAGGATGCTTCTCTGATCCCTGTACA 2824  
DB GGAACAGAGTGGTTTACANGTCTCGACCTTCAGGATGCTTCTGATCCCTGTACA 5078  
QY 2825 TCCTGACTCTCAATTTCTGTTGCTTTGAAGATCTTCAAACCCCAACATCTCAACTCAC 2884  
DB TCCTGACTCTCAATTTCTGTTGCTTTGAAGATCTTCAAACCCCAACATCTCAACTCAC 5138  
QY 2885 CTGAGCTRTTTACCCCAAGGCTTCAGGATAGYCCCCATCTATTTGGCCAGGCAATTAGC 2944  
DB CTGAGCTRTTTACCCCAAGGCTTCAGGATAGTCCCCATCTATTTGGCCAGGCAATTAGC 5198  
QY 2945 CCAAGACTTGAGYCAATYMTACATCTGGACACTCTGTGCTTCRTGTAGTGGATGATTT 3004  
DB CCAAGACTTGAGYCAATCTCATACCTGGACA--CTGTGCTTCRTGTAGTGGATGATTT 5256  
QY 3005 ACTTTTGGYGCYRFTCAGAAACCTGTGCCATCAAGCACCCCAAGGCTCTTMAATTT 3064  
DB ACTTTTGGYGCYRFTCAGAAACCTGTGCCATCAAGCACCCCAAGGCTCTTMAATTT 5316  
QY 3065 CCTCCYACCTGTGGCTCAWGGTTTCCAAAGSABRGCTCARCTGCTCACAGCAGGT 3124  
DB CCTCCYACCTGTGGCTCAWGGTTTCCAAAGSABRGCTCARCTGCTCACAGCAGGT 5376  
QY 3125 TAAATACCTTAGRCTAAATATCAAAAGGCACCAAGGCCCTCAGTGAAGGAAYRATCCA 3184  
DB TAAATACCTTAGRCTAAATATCAAAAGGCACCAAGGCCCTCAGTGAAGGAAYRATCCA 5377  
QY 3185 GCCTATACCTGCTTATCTCATCYCAAAACCTTAAAGCAACTAAGRRTTCTCTGGCT 3244  
DB GCCTATACCTGCTTATCTCATCYCAAAACCTTAAAGCAACTAAGRRTTCTCTGGCT 5492  
QY 3245 AAYAGGYTCTGCCGAATGATTTCCCGAGTGTGGCBAATAGCCAGGYCATTAWATA 3304  
DB AAYAGGYTCTGCCGAATGATTTCCCGAGTGTGGCBAATAGCCAGGYCATTAWATA 5493  
QY 3305 CASTAATTAAGGAACCTCAGAAAGCAATPACCCATTTARTAGATGGAYMCTGAAGYMR 3364  
DB CASTAATTAAGGAACCTCAGAAAGCAATPACCCATTTARTAGATGGAYMCTGAAGYMR 5551  
QY 3552 CACTAATTAAGGAACCTCAGAAAGCAATPACCCATTTAGTAGATGGACACTGAAGTAG 5611  
DB CACTAATTAAGGAACCTCAGAAAGCAATPACCCATTTAGTAGATGGACACTGAAGTAG 5611  
QY 3365 AAGTGGCTTTCCAGGCCCTTAAAGAGGCTTTAAACCCCAAGYCCAGGTGTTAAGYTTGCC 3424  
DB AAGTGGCTTTCCAGGCCCTTAAAGAGGCTTTAAACCCCAAGYCCAGGTGTTAAGYTTGCC 5657  
QY 3425 AACRGGGCAGACTTTTSTYATATVATCAGAGAAACAGAAAYAGCTCTRGAGTCC 3484  
DB AACRGGGCAGACTTTTSTYATATVATCAGAGAAACAGAAAYAGCTCTRGAGTCC 5658  
QY 3485 TTACACAGTCCRAGGGYAGCTTGGCAACCYRTGGCRVACCTGASTAAGGAAAYTGATG 3544  
DB TTACACAGTCCRAGGGYAGCTTGGCAACCYRTGGCRVACCTGASTAAGGAAAYTGATG 5717  
QY 3545 TAGTGGCAAGGTTGRCYTCATTTGTTAYVGGGTAGTGGCAGTAGCAGTYKTAGTAT 3604  
DB TAGTGGCAAGGTTGRCYTCATTTGTTAYVGGGTAGTGGCAGTAGCAGTYKTAGTAT 5777  
QY 3604 ARAACAAARAAACAGGCCATTACCTGTRARAARAACTGGGCAACTGATTTTCCACCAAGCCC 4743

QY 3605 CTGAAGCAGTTAAATTAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGA 3664  
DB CTGAAGCAGTTAAATTAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGA 5896  
QY 3665 AYRGATATCTCACTGCTAAAGGAGACTTGTGGCTGTGAGACAACYGTGTTACTTAATPTC 3724  
DB ATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTGAGACAACYGTGTTACTTAATPTC 5956  
QY 3725 AGGCTCTATTACTTTGAARGGCCAGTGTGACCTGTGACCTGTGCAACTCTTAACCCAG 3784  
DB AGGCTCTATTACTTTGAARGGCCAGTGTGACCTGTGACCTGTGCAACTCTTAACCCAG 6016  
QY 3785 YCNCAATTTCTTCCAGACAATGAAGAAAGATARAAYATAACTGTCAAACAARTAAATTTCTC 3844  
DB CCACATTTCTTCCAGACAATGAAGAAAGATARAAYATAACTGTCAAACAARTAAATTTCTC 6076  
QY 3845 AAACCTATGCACTCGAGGGGACCTTCTTAGARGTTTCYTTGACTGATCCYGAACCTTCAAC 3904  
DB AAACCTATGCACTCGAGGGGACCTTCTTAGARGTTTCYTTGACTGATCCYGAACCTTCAAC 6135  
QY 3905 TTGTATATCTGATGGAAGTTCTTTGTAGAAAAGAGCTTCGAAAAGYGGGTATGAGTG 3964  
DB TTGTATATCTGATGGAAGTTCTTTGTAGAAAAGAGCTTCGAAAAGYGGGTATGAGTG 6195  
QY 3965 GTCACTGATATGAATATTTGAAGTAATCCCTCTCACTCCAGGAACTAGTGTGCTVAGCTR 4024  
DB GTCACTGATATGAATATTTGAAGTAATCCCTCTCACTCCAGGAACTAGTGTGCTVAGCTR 6255  
QY 4025 GCAGAACTAATAGCCCTCAYTKGGGCACCTAGAAATTTAGGAGAAAGGAAAGGAAATATA 4084  
DB GCAGAACTAATAGCCCTCAYTKGGGCACCTAGAAATTTAGGAGAAAGGAAAGGAAATATA 6315  
QY 4085 TATCAGACTCTATATATGCTTACTAGTCTTCCATGCTCCATGMRGAATATGSRAGAA 4144  
DB TATCAGACTCTATATATGCTTACTAGTCTTCCATGCTCCATGMRGAATATGSRAGAA 6375  
QY 4145 AGGAAATCTCTAACTTCYAGAGGAACACTTATCAGACATCAGGAGCCATTTAGGARATTA 4204  
DB AGGAAATCTCTAACTTCYAGAGGAACACTTATCAGACATCAGGAGCCATTTAGGARATTA 6435  
QY 4205 TTATYGGCWGTACAGAAACCTTARAGAGTGGMAGTCTTACACTGCGGGGTCAATCANAA 4264  
DB TTATYGGCWGTACAGAAACCTTARAGAGTGGMAGTCTTACACTGCGGGGTCAATCANAA 6495  
QY 4265 GGAAGRAAGAGGAAATASAGRGAATGCGCAAGCAKATATTTGAAGCAAAAGAGCTGCA 4324  
DB GGAAGRAAGAGGAAATASAGRGAATGCGCAAGCAKATATTTGAAGCAAAAGAGCTGCA 6555  
QY 4325 AGGAGGACCTCCATTTAGAAATGCTTATATAAACCAACCCCTAGTATAGGTTAATCCCTC 4384  
DB AGGAGGACCTCCATTTAGAAATGCTTATATAAACCAACCCCTAGTATAGGTTAATCCCTC 6615  
QY 4385 CGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAAATGGGAACTCTCAGAGG--CAG 4443  
DB CGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAAATGGGAACTCTCAGAGG--CAG 6675  
QY 4444 TTTTCTCCCTCGGAGCGGTTAGCACCTGGAAGAGGAAATACTTTTTCCTGCTGCAACTAT 4503  
DB TTTTCTCCCTCGGAGCGGTTAGCACCTGGAAGAGGAAATACTTTTTCCTGCTGCAACTAT 6735  
QY 4504 CCAATGGAAATTACTTAAACCCCTTCAACACCTTTTCACTTAGGCATCGATGACCCCA 4563  
DB CCAATGGAAATTACTTAAACCCCTTCAACACCTTTTCACTTAGGCATCGATGACCCCA 6795  
QY 4564 TCARATGGCCAAATCATTTATTTA CTGGACAGGCCCTTTTCAAAA CTATCAAGCARATATK 4623  
DB TCARATGGCCAAATCATTTATTTA CTGGACAGGCCCTTTTCAAAA CTATCAAGCARATATK 6855  
QY 4624 CAGGCTCTGTGAATGTGCCARARAATAATCCCTGCTTATCCGCAAGCTCTTTCAGG 4683  
DB CAGGCTCTGTGAATGTGCCARARAATAATCCCTGCTTATCCGCAAGCTCTTTCAGG 6915  
QY 4684 ARAACAAARAAACAGGCCATTACCTGTRARAARAACTGGGCAACTGATTTTCCACCAAGCCC 4743



Qy	6904	AGCCATGAGATGCGCTGGATTCTCCCTTCTTAGGACTCTAGCAGCTATAATATGCTA	6963
Db	9135	AGCCAAATGGATGCGCTGGATTCTCCCTTCTTAGGACTCTAGCAGCTATAATATGCTA	9194
Qy	6964	CTCCTCTTTGGACCGCTGTATCTTTTACCTCTCTTGTAACTTGTCTTCCAGAAATCGAA	7023
Db	9195	CTCCTCTTTGGACCGCTGTATCTTTTAACTCTCTTGTAACTTGTCTTCCAGAAATCGAA	9254
Qy	7024	GCTGTRAAACTACAAATGGAGCCCAAGATGAGTCCAGACTAAGATCTACCGCAGACCC	7083
Db	9255	GCTGTRAAACTACAAATGGAGCCCAAGATGAGTCCAGACTAAGATCTACCGCAGACCC	9314
Qy	7084	CTGACCGGCTGTAGCCCAAGATCTGTATGTTAATGACATCAAGGACCCCTCTCTGAG	7143
Db	9315	CTGACCGGCTGTAGCCCAAGATCTGTATGTTAATGACATCAAGGACCCCTCTCTGAG	9374
Qy	7144	GAATCTCAGTGTGAACCTCTACTAGCCGCCAATTCAGCAGGAAGAGTTAGAGCGGT	7203
Db	9375	GAATCTCAGTGTGAACCTCTACTAGCCGCCAATTCAGCAGGAAGAGTTAGAGCGGT	9434
Qy	7204	SETCGGCCAACCTCCCAACAGCACTTAGGTTTCTGTTGAGATGGGGACTGAGAGAC	7263
Db	9435	SETCGGCCAACCTCCCAACAGCACTTAGGTTTCTGTTGAGATGGGGACTGAGAGAC	9494
Qy	7264	AGGACTAGCTGGATTCTTAGGCTGATTAAGAAATCCCTTAAGCCCTAGGTGGGAAGTCAAC	7323
Db	9495	AGGACTAGCTGGATTCTTAGGCTGATTAAGAAATCCCTTAAGCCCTAGGTGGGAAGTCAAC	9554
Qy	7324	ACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCAACCTGACCAATCAGAGAGCTC	7383
Db	9555	ACATCCACCTTTAAACACGGGGCTTGCAACTTAGYTCAACCTGACCAATCAGAGAGCTC	9614
Qy	7384	ACTAAATGCTAATTAGGCAAGACAGGAGTAAAGAAATAGCCAAATCATYTTATGCGTG	7443
Db	9615	ACTAAATGCTAATTAGGCAAGACAGGAGTAAAGAAATAGCCAAATCATYTTATGCGTG	9674
Qy	7444	AGACACAGCAGGAGGACAAATGATCGGATATAAACCCAGTCTTCGAGCCGCGCAACGG	7503
Db	9675	AGACACAGCAGGAGGACAAATGATCGGATATAAACCCAGTCTTCGAGCCGCGCAACGG	9734
Qy	7504	CAACCCCTTTGGTCCCTCCCTTTGTATGGGAGCTCTGTTTTTCATGCTATTTCACTCT	7563
Db	9735	CAACCCCTTTGGTCCCTCCCTTTGTATGGGAGCTCTGTTTTTCATGCTATTTCACTCT	9794
Qy	7564	ATTAATCTTGCACCTGCR 7582	
Db	9795	ATTAATCTTGCACCTGCA 9813	
RESULT 6			
AY101585		10222 bp	DNA linear PRI 11-FEB-2004
LOCUS			
DEFINITION	Homo sapiens isolate 79 endogenous retrovirus HERV-W, ERVW1 locus, allele B, complete sequence.		
ACCESSION	AY101585		
VERSION	AY101585.1	GI:37544405	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 10222) Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.		
TITLE	The endogenous retroviral locus ERVW1 is a bona fide gene involved in hominoid placental physiology		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)		
PUBLISHED	14757826		
REFERENCE	2 (bases 1 to 10222)		
AUTHORS	Mallet, F., Bouton, O. and Oriol, G.		
TITLE	Direct Submission		





QY 3185 GCCTATCTGGCTTATCTCTCATCYCAAAACCCCTAAAGCAACTAAAGRRRTTCTTGGCT 3244  
DB 5433 GCCTATCTGGCTTATCTCTCATCCCAAAACCCCTAAAGCAACTAAAGGGGATTCCTTTGGCT 5492  
QY 3245 AAYAGGYTTCTCCGGAATGATTTCCCGAGTWTGGRAAATAGCCAGGYCATTAATA 3304  
DB 5493 AATAGGTTTCTCCGGAATGATTTCCCGAGTWTGGRAAATAGCCAGGYCATTAATA 5551  
QY 3305 CASTAATTAAGGAATCTCAGAAAGCAATACCATTARTTAAGATGGAYAMCTGAAGYMR 3364  
DB 5552 CACTAATTAAGGAATCTCAGAAAGCAATACCATTARTTAAGATGGAYAMCTGAAGYMR 5611  
QY 3365 AAGTGGCTTTCAGGCCCCCTAAGAGGCTTAAACCCAGYCCAGGTTTAAAGYTTGCC 3424  
DB 5612 AAGTGGCTTTCAGGCCCCCTAAGAGGCTTAAACCCAGYCCAGGTTTAAAGYTTGCC 5657  
QY 3425 AACRGGGCAAGACTTTTSTTYATAYRTCAAGAAAAACAGAAAYAGCTCTRGAGTCC 3484  
DB 5658 AACAGGGCAAGACTTTTCTCATATGTACAG-AAAAAACAGAAATAGCTCTAGGAGTCC 5716  
QY 3485 TTACACAGTTCRAGGGAYGAGCTTGCACCCYRTGCRYACCTGASTAAGGAAAYTGATG 3544  
DB 5717 TTACACAGATCCAGGGATGAGCTTGCACCCYRTGCRYACCTGASTAAGGAAAYTGATG 5776  
QY 3545 TAGTGGCAAGGTTGRCYTCATTTGTTTAYGGGTAGTGGTGGAGTACAGGAGTCTAGTAT 3604  
DB 5777 TAGTGGCAAGGTTGAGCTCATTTGTTTACGGGTAGTGGTGGAGTACAGGAGTCTAGTAT 5836  
QY 3605 CTGAGGAGTAAATAATAACAGGGAGAGATCTTACTGTGTGGACATCTCATGAKGTGA 3664  
DB 5837 CTGAGGAGTAAATAATAACAGGGAGAGATCTTACTGTGTGGACATCTCATGAGTGA 5896  
QY 3665 AYRGATATCTACTGTCTAAGGAGACTTGTGGCTGTGCAGCACTGTTTACTTAAATRTC 3724  
DB 5897 ATGGCATATCTACTGTCTAAGGAGACTTGTGGCTGTGCAGCACTGTTTACTTAAATGTC 5956  
QY 3725 AGGCTATATCTTGAAGGCGAGTGTGCTGACACTGTGCACTGTGGAACCTTAAACCGAG 3784  
DB 5957 AGGCTATATCTTGAAGGCGAGTGTGCTGACACTGTGCACTGTGGAACCTTAAACCGAG 6016  
QY 3785 YNCATATCTTCCAGACATGAAGAAAGATARAAYATAACTGTCAACAARTAAATTCCTC 3844  
DB 6017 CCACATTTCTCCAGACATGAAGAAAGATARAAYATAACTGTCAACAARTAAATTCCTC 6076  
QY 3845 AAACCTATGCCACTCGAGGGGACCTTCTTAGAGTTCCTTGTGACTGATCCCGACC-TCAAC 6135  
DB 6077 AAACCTATGCCACTCGAGGGGACCTTCTTAGAGTTCCTTGTGACTGATCCCGACC-TCAAC 6135  
QY 3905 TTGTATATGATGGAAGTTCTTTGTAGAAAAGGACCTGAAAAGYGGGTATGCAATG 3964  
DB 6136 TTGTATATGATGGAAGTTCTTTGTAGAAAAGGACCTTGTGAAAAGTGGGGTATGCAATG 6195  
QY 3965 GTCAAGTAAATGGAATAVTTGAAAGTAAATCCCTTCACTCCAGGACCTAGTCTAGCTC 4024  
DB 6196 GTCAAGTAAATGGAATAVTTGAAAGTAAATCCCTTCACTCCAGGACCTAGTCTAGCTC 6255  
QY 4025 GCAGAACTAATAGCCCTCTTCTGAGGCACTAGAAATAGGAGAGAAAAAGGGYAAATATA 4084  
DB 6256 GCAGAACTAATAGCCCTCTTCTGAGGCACTAGAAATAGGAGAGAAAAAGGGYAAATATA 6315  
QY 4085 TATACAGATCTTARTATGCTTACCTAGTCTGTCATGTCAGGAGTATGAGAGAA 4144  
DB 6316 TATACAGATCTTARTATGCTTACCTAGTCTGTCATGTCAGGAGTATGAGAGAA 6375  
QY 4145 AGGAAATCTTAACTTCTCAGAGAACACCTATCAMAATCAGAGGCACTTAGGAGATTA 4204  
DB 6376 AGGAAATCTTAACTTCTCAGAGAACACCTATCAMAATCAGAGGCACTTAGGAGATTA 6435  
QY 4205 TTATGCGGTGACAGAAACCTTARAGAGGTGGMAGTCTTACACTGCGGGGTCTCANAAA 4264  
DB 6436 TTATGCGGTGACAGAAACCTTARAGAGGTGGMAGTCTTACACTGCGGGGTCTCANAAA 6495

QY 4265 GGAAGRAAAGGGAATASRAAGRAAYTGCCAAAGCAKATATTTGAAGCAAAAGAGCTGCA 4324  
DB 6496 GGAAGRAAAGGGAATASRAAGRAAYTGCCAAAGCAKATATTTGAAGCAAAAGAGCTGCA 6555  
QY 4325 AGGAGGACCCCTCCATTAGAAATGTTTATAAACTTCCCTTTAGTATAGGGTAATCCCTTC 4384  
DB 6556 AGGAGGACCCCTCCATTAGAAATGTTTATAAACTTCCCTTTAGTATAGGGTAATCCCTTC 6615  
QY 4385 CGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAAATGGGNAACCTCAGCAGG-CAG 4443  
DB 6616 CGGAAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAAATGGGNAACCTCAGCAGGAGCAG 6675  
QY 4444 TTTTCTCCCTCGGAGCGTTAGCACCTGAAGAGGGAATACTTTTGGCTGCAACTAT 4503  
DB 6676 TTTTCTCCCTCGGAGCGCTAGCCACTGAAGAGGGAATACTTTTGGCTGCAACTAT 6735  
QY 4504 CCAATGGAAATTACTTTAAACCCCTTCATCAAACTTTTCACTTTAGGCATCGATAGCACCA 4563  
DB 6736 CCAATGGAAATTACTTTAAACCCCTTCATCAAACTTTTCACTTTAGGCATCGATAGCACCA 6795  
QY 4564 TCARATGGCCAAATCATTTATTTA CTGGAACAGGCGCTTTTCAAACTATCAAGCARATAT 4623  
DB 6796 TCARATGGCCAAATCATTTATTTA CTGGAACAGGCGCTTTTCAAACTATCAAGCARATAT 6855  
QY 4624 CAGGCGCTGTGAATGTGCGARARAAATAATCCCTGCTATCCCAAGCTCTTTCAGG 4683  
DB 6856 CAGGCGCTGTGAATGTGCGARARAAATAATCCCTGCTATCCCAAGCTCTTTCAGG 6915  
QY 4684 ARAACAAARAAACAGGCCATTAATCCCTGARAARAACTTGGCAACTGTATTTTACCCACAAGCCC 4743  
DB 6916 ARAACAAARAAACAGGCCATTAATCCCTGARAARAACTTGGCAACTGTATTTTACCCACAAGCCC 6975  
QY 4744 AAACCTCAGGATTTTCACTATCTAGTCTGCGGTATATATCTTTCACCGGTTGGGAGAG 4803  
DB 6976 AAACCTCAGGATTTTCACTATCTAGTCTGCGGTATATATCTTTCACCGGTTGGGAGAG 7035  
QY 4804 GCCTTCCCTCTGAGGAGAAAGCCCAAGAGGTAATAAAGGCACTAGTTTCACTGAAATA 4863  
DB 7036 GCCTTCCCTCTGAGGAGAAAGCCCAAGAGGTAATAAAGGCACTAGTTTCACTGAAATA 7095  
QY 4864 ATTCCAGATTCGGACTTTCCCGAGGCTTACAGAGTGACAAATAGCCCTTCTTTCAGGCCC 4923  
DB 7096 ATTCCAGATTCGGACTTTCCCGAGGCTTACAGAGTGACAAATAGCCCTTCTTTCAGGCCC 7155  
QY 4924 ACAGTAAACCCAGGAGTATCCAGCGGTATAGGTATAGTATCACTTACCTGCGGCTGA 4983  
DB 7156 ACAGTAAACCCAGGAGTATCCAGCGGTATAGGTATAGTATCACTTACCTGCGGCTGA 7215  
QY 4984 AGGCCACAGTCTCAGGAGAGGTCGAGAAATGAATGAAAYACTCAAGGACATCTAAAA 5043  
DB 7216 AGGCCACAGTCTCAGGAGAGGTCGAGAAATGAATGAAAYACTCAAGGACATCTAAAA 7275  
QY 5044 AAGCAAAACCCAGGAAACCCCTCATCTGCGCTGTCTGTGTGCTTATAGCCTTAAAAAGA 5103  
DB 7276 AAGCAAAACCCAGGAAACCCCTCATCTGCGCTGTCTGTGTGCTTATAGCCTTAAAAAGA 7335  
QY 5104 ATCTGCAATTTTCCCAAAAGAGGACTTATAGCCCATAGCAAAATGCTGTATGAGAGGCC 5163  
DB 7336 ATCTGCAATTTTCCCAAAAGAGGACTTATAGCCCATAGCAAAATGCTGTATGAGAGGCC 7395  
QY 5164 TTCTAAACCAATGACTTGTGCTTGCACCAAGAGGCACTTGTAGTGTGAGACATCACCT 5223  
DB 7396 TTCTAAACCAATGACTTGTGCTTGCACCAAGAGGCACTTGTAGTGTGAGACATCACCT 7455  
QY 5224 CCTTAGCCAAATATCAAAAGTCTTTAAAACTTCAAGAAACCTTATCCCTGAGAGAGAG 5283  
DB 7456 CCTTAGCCAAATATCAAAAGTCTTTAAAACTTCAAGAAACCTTATCCCTGAGAGAGAG 7515  
QY 5284 GAAAGAACTATTTCCACCCWGTGATAGGTATAGTCAAGTCCCTTCTCTTAATCCC 5343  
DB 7516 GAAAGAACTATTTCCACCCWGTGATAGGTATAGTCAAGTCCCTTCTCTTAATCCC 7575  
QY 5344 CATCCCTAGATACATCTCTGGGAGGAGCCCTACCCAGTCTATTTTATVTTACCCCACTGCGG 5403



Db 7576 CATCCCTAGATACATCTCGGAAGGACCTACCCAGTCATTTTATCTACCCCACTGCGG 7635  
Qy 5404 TTAAGTGGCTGGAGTGGAGCTTTGGATACATCACACTTGGAGCAAAATCCTGATACTGC 5463  
Db 7636 TTAAGTGGCTGGAGTGGAGCTTTGGATACATCACACTTGGAGCAAAATCCTGATACTGC 7695  
Qy 5464 CAAAGGAACCTGAAATCCAGGAGACAAACGCTAGCTATTCTCTGAACTCTAGAGGATT 5523  
Db 7696 CAAAGGAACCTGAAATCCAGGAGACAAACGCTAGCTATTCTCTGAACTCTAGAGGATT 7755  
Qy 5524 TGGCCTGCTCTTCAACCAACACGAGGAGAAAGTAACTAAATCATAAATCCCCCATG 5583  
Db 7756 TGGCCTGCTCTTCAACCAACACGAGGAGAAAGTAACTAAATCATAAAT-CCCCATG 7814  
Qy 5584 GSCCTCCCTTATCATATTTTCTCTTASTGTTSTTTTACCTCTTCACTCTCACTGCA 5643  
Db 7815 GSCCTCCCTTATCATATTTTCTCTTACTGTCTTTTACCTCTTCACTCTCACTGCA 7874  
Qy 5644 CCCCTCCATGCGCTGTATGACCACTAGCTAGCTCCCTTACCMAGAGTTCTATGGAGATG 5703  
Db 7875 CCCCTCCATGCGCTGTATGACCACTAGCTAGCTCCCTTACCMAGAGTTCTATGGAGATG 7934  
Qy 5704 CAGCGTCCCGGAAATATGATGCCCACTGTATAGGAGTCTTTSTAAGGGAACCCCAACC 5763  
Db 7935 CAGCGTCCCGGAAATATGATGCCCACTGTATAGGAGTCTTTSTAAGGGAACCCCAACC 7994  
Qy 5764 TTCACTGCCCAACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCACTGCAT 5823  
Db 7995 TTCACTGCCCAACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCACTGCAT 8054  
Qy 5824 GCAATACTCATTTTGGACAGGAAAAATGATTAATCTTAGTGTGCTGAGAGACTTGA 5883  
Db 8055 GCAATACTCATTTTGGACAGGAAAAATGATTAATCTTAGTGTGCTGAGAGACTTGA 8114  
Qy 5884 GTCACTGTCTGTGGAGTCTTACCCCAACCTGATGTCATGCGGGGTGAGTTCAA 5943  
Db 8115 GTCACTGTCTGTGGAGTCTTACCCCAACCTGATGTCATGCGGGGTGAGTTCAA 8174  
Qy 5944 GATCAGGCAAGAGAAAAATGTAAAGAGTAATCTCCCAACTCAGCGGGGTACATGGC 6003  
Db 8175 GATCAGGCAAGAGAAAAATGTAAAGAGTAATCTCCCAACTCAGCGGGGTACATGGC 8234  
Qy 6004 ACCTCTAGCCCTACAAAGACTAGATCTCTCAAACTACATAGAAACCCCTCGATACCCAT 6063  
Db 8235 ACCTCTAGCCCTACAAAGACTAGATCTCTCAAACTACATAGAAACCCCTCGATACCCAT 8294  
Qy 6064 ACTCGCCTGTAAGCCTATTTAATACCAACCTCAGCTGGGCTCCATGAGTCTCGGCCAA 6123  
Db 8295 ACTCGCCTGTAAGCCTATTTAATACCAACCTCAGCTGGGCTCCATGAGTCTCGGCCAA 8354  
Qy 6124 AACCCTACTAACTGTTGGATATGCCCTCCCTGAACTTCARGGCATATGTTTCAATCCCT 6183  
Db 8355 AACCCTACTAACTGTTGGATATGCCCTCCCTGAACTTCARGGCATATGTTTCAATCCCT 8414  
Qy 6184 GTACCTGAACAATGGAAACAACTTCAGCAAGAAATAAACAACCTTCGTTTTAGTAGGA 6243  
Db 8415 GTACCTGAACAATGGAAACAACTTCAGCAAGAAATAAACAACCTTCGTTTTAGTAGGA 8474  
Qy 6244 CCTCTGTTTCCAAATSTGGAATAACCCATACCTCAAACTCAGCTGTGTAATAATTTAGC 6303  
Db 8475 CCTCTGTTTCCAAATSTGGAATAACCCATACCTCAAACTCAGCTGTGTAATAATTTAGC 8534  
Qy 6304 AATACTACATACAAACCAACTCCCAATGATCAGGTGGGTAACTCTCCCAACCAATA 6363  
Db 8535 AATACTACATACAAACCAACTCCCAATGATCAGGTGGGTAACTCTCCCAACCAATA 8594  
Qy 6364 GTCTGCTACCTCAGGAATAATTTTGTGTGTGTGCTGAGTCTCAGCCTATCGTTGTTGAAT 6423  
Db 8595 GTCTGCTACCTCAGGAATAATTTTGTGTGTGTGCTGAGTCTCAGCCTATCGTTGTTGAAT 8654  
Qy 6424 GGCCTTTCAGATCTATGCTTCTCTCATTTCTTAGTGGCCCTCCTATCTACACT 6483

Db 8655 GGCCTTTCAGATCTATGCTTCTCTCATTTCTTAGTGCCCTATGACCATCTACACT 8714  
Qy 6484 GAAACAAGATTTATACAGTTATGTATATCTAAGCCCGCAACAAAGAGTACCAATCTT 6543  
Db 8715 GAAACAAGATTTATACAGTTATGTATCTAAGCCCGCAACAAAGAGTACCAATCTT 8774  
Qy 6544 CCTTTTGTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCAATGGCGGTATCACA 6603  
Db 8775 CCTTTTGTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCAATGGCGGTATCACA 8834  
Qy 6604 ACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGACATGGAAACGGTC 6663  
Db 8835 ACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGACATGGAAACGGTC 8894  
Qy 6664 GCGACTCCCTGCTCACTTGTCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTCRA 6723  
Db 8895 GCGACTCCCTGCTCACTTGTCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTCRA 8954  
Qy 6724 AATCGAAGACTTTAGACTTGTACCGCTGABAGAGGGGGAACCTGTTTATTTTAGGG 6783  
Db 8955 AATCGAAGACTTTAGACTTGTACCGCTGABAGAGGGGGAACCTGTTTATTTTAGGG 9014  
Qy 6784 GAAGAAATGCTGTTTATTTATGTTTAACTCCGGAATCGTCACTGAGAAAGTTAAAGAAAT 6843  
Db 9015 GAAGAAATGCTGTTTATTTATGTTTAACTCCGGAATCGTCACTGAGAAAGTTAAAGAAAT 9074  
Qy 6844 CSAGATCGAATCAACGTAAGCAGAGAGACTTTCGAAACAACCTGGACCTCGGGGCTCTCT 6903  
Db 9075 CGAGATCGAATCAACGTAAGCAGAGAGACTTTCGAAACAACCTGGACCTCGGGGCTCTCT 9134  
Qy 6904 AGCCRAATGATGCCCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATATATTTGCTA 6963  
Db 9135 AGCCRAATGATGCCCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATATATTTGCTA 9194  
Qy 6964 CTCCTCTTTGGACCTCTGATCTTTTACCTCTGTTTAACTTTGTCTCTTCCAGAAATCGAA 7023  
Db 9195 CTCCTCTTTGGACCTCTGATCTTTTAACTCTTGTGTTTAACTTTGTCTCTTCCAGAAATCGAA 9254  
Qy 7024 GCTGTRAAACTCAAAATGGAGCCCAAGATGCACTCAAGACTAAAGATCTACCGCAGACCC 7083  
Db 9255 GCTGTRAAACTCAAAATGGAGCCCAAGATGCACTCAAGACTAAAGATCTACCGCAGACCC 9314  
Qy 7084 CTGGAACCGGCTGTATAGCCCAAGATCTGATGTTTAAATGATCAAAAGGCAACCCCTCTGAG 7143  
Db 9315 CTGGAACCGGCTGTATAGCCCAAGATCTGATGTTTAAATGATCAAAAGGCAACCCCTCTGAG 9374  
Qy 7144 GAAATCTCAGCTGCAACCACTCTACTACGCCCCCAATTCAGCAGGAGCACTTAGACGGT 7203  
Db 9375 GAAATCTCAGCTGCAACCACTCTACTACGCCCCCAATTCAGCAGGAGCACTTAGACGGT 9434  
Qy 7204 SGTGGGCCAAACCTCCCAACAGCACTTAGGTTTCTCTGTTGAGATGGGGGACTGAGAGAC 7263  
Db 9435 CGTGGGCCAAACCTCCCAACAGCACTTAGGTTTCTCTGTTGAGATGGGGGACTGAGAGAC 9494  
Qy 7264 AGGACTAGCTGGAATTTCTAGGCTGATTAAGAAATCCYTAAGCCTAGSTGGGAAGGTGACC 7323  
Db 9495 AGGACTAGCTGGAATTTCTAGGCTGATTAAGAAATCCYTAAGCCTAGSTGGGAAGGTGACC 9554  
Qy 7324 ACATCCACTTTTAAACACGGGCTTGCACACTTAGTGTACACCTGACCAATTCAGAGAGCTC 7383  
Db 9555 ACATCCACTTTTAAACACGGGCTTGCACACTTAGTGTACACCTGACCAATTCAGAGAGCTC 9614  
Qy 7384 ACTAAAATGCTAATTTAGGCAAGACAGGAGTAAAGAAATAGCCAAATCATYTTATTCMTG 7443  
Db 9615 ACTAAAATGCTAATTTAGGCAAGACAGGAGTAAAGAAATAGCCAAATCATYTTATTCMTG 9674  
Qy 7444 AGAGCAGCAGCAGGAGGCAATGATCGGGATATAAACCCAAAGTTCGAGCCGGCAACGG 7503  
Db 9675 AGAGCAGCAGCAGGAGGCAATGATCGGGATATAAACCCAAAGTTCGAGCCGGCAACGG 9734  
Qy 7504 CAAACCCCTTTGGGTCCTCTCCCTTGTATGAGGAGCTCTGTTTTCAATGCTATTTCACTCT 7563  
Db 9735 CAAACCCCTTTGGGTCCTCTCCCTTGTATGAGGAGCTCTGTTTTCAATGCTATTTCACTCT 9794





QY 4444 TTTTCTCCCTCGGACGGTTAGCCACTGAAGAAGGAAATACATTTTGGCTGCAACTAT 4503  
DB |||||  
DB 34743 TTTTCTCCCTCGGACGGTCTAGCCACTGAAGAAGGAAATACATTTTGGCTGCAACTAT 34802  
QY 4504 CCAATGGAAATTTACCTTAAACCCCTTTTCAACACCTTTTCACTTTAGGCATCGATAGCACCCA 4563  
DB |||||  
DB 34803 CCAATGGAAATTTACCTTAAACCCCTTTTCAACACCTTTTCACTTTAGGCATCGATAGCACCCA 34862  
QY 4564 TCARATGGCAATCATTTATTTACTTGGACAGGCCCTTTTCAAAAACCTATCAAGCARATKAT 4623  
DB |||||  
DB 34863 TCAGATGGCAATCATTTATTTACTTGGACAGGCCCTTTTCAAAAACCTATCAAGCARATKAT 34922  
QY 4624 CAGGSCCTGTGAAGTGTGCGAGAGAAATTAATCCCTGCTTATGCCAAGCTTCTTTCAGG 4683  
DB |||||  
DB 34923 CAGGSCCTGTGAAGTGTGCGAGAGAAATTAATCCCTGCTTATGCCAAGCTTCTTTCAGG 34982  
QY 4684 AFAACAAARAAACAGGCCATTTACCTTGRARAARACTTGGCAACTGATTTTACCCACAAGGCC 4743  
DB |||||  
DB 34983 AGAACAAAGNAACAGGCCATTTACCTTGGAGAGACTTGGCAACTGATTTTACCCACAAGGCC 35042  
QY 4744 AAACCTCAGGGATTTTCACTATCTACTAGTCTGGGTARATATCTTTTCAAGGGTTGGCARAG 4803  
DB |||||  
DB 35043 AAACCTCAGGGATTTTCACTATCTACTAGTCTGGGTARATATCTTTTCAAGGGTTGGCARAG 35102  
QY 4804 GCCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAAGGCACCTAGTTTCATGAATA 4863  
DB |||||  
DB 35103 GCCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAAGGCACCTAGTTTCATGAATA 35162  
QY 4864 ATTTCCAGATTCGGACTTCCCGAGGCTTTACAGAGTGACAATAGGCCCTTTCAGAGGCC 4923  
DB |||||  
DB 35163 ATTTCCAGATTCGGACTTCCCGAGGCTTTACAGAGTGACAATAGGCCCTTTCAGAGGCC 35222  
QY 4924 ACAGTAACCCAGGGAGTATCCAGGGCTTGTAGTATACGATATCACTTACCTGCGCCTGA 4983  
DB |||||  
DB 35223 ACAGTAACCCAGGGAGTATCCAGGGCTTGTAGTATACGATATCACTTACCTGCGCCTGA 35282  
QY 4984 AGGCCACAGTCTCAGGGAGGTTCAGAGAAATGAATGAAGAACTCAAGGACATCTAAA 5043  
DB |||||  
DB 35283 AGGCCACAGTCTCAGGGAGGTTCAGAGAAATGAATGAAGAACTCAAGGACATCTAAA 35342  
QY 5044 AAGCAAAACCCAGGAAACCCACCTCAGTGGCTGTCTGTCTGCTTATAGCCCTTAAAGAA 5103  
DB |||||  
DB 35343 AAGCAAAACCCAGGAAACCCACCTCAGTGGCTGTCTGTCTGCTTATAGCCCTTAAAGAA 35402  
QY 5104 ATCTGCACTTTCCCAAAAGCAGGACTTACGCCATACGAAATGCTGTATGGAAGGCC 5163  
DB |||||  
DB 35403 ATCTGCACTTTCCCAAAAGCAGGACTTACGCCATACGAAATGCTGTATGGAAGGCC 35462  
QY 5164 TTCTAATCAATGACCTTGTGCTTGAACCAAGCAGCAGCACTTACCTGAGACATCACT 5223  
DB |||||  
DB 35463 TTCTAATCAATGACCTTGTGCTTGAACCAAGCAGCAGCACTTACCTGAGACATCACT 35522  
QY 5224 CCTTAGCCAAATATCAACAAATTTCTTAAACATTTACAGGAACCTTATCCCTGAGAGAGG 5283  
DB |||||  
DB 35523 CCTTAGCCAAATATCAACAAATTTCTTAAACATTTACAGGAACCTTATCCCTGAGAGAGG 35582  
QY 5284 GAAAGAACTATTTCAACCCCMGTGACATGGTATAGTCAAGTCCCTTCTTAATCCC 5343  
DB |||||  
DB 35583 GAAAGAACTATTTCAACCCCMGTGACATGGTATAGTCAAGTCCCTTCTTAATCCC 35642  
QY 5344 CATCCCTAGATACATCTTGGAGAGGACCTTACCCAGTCAATTTATTTATCCCAACTGCGG 5403  
DB |||||  
DB 35643 CATCCCTAGATACATCTTGGAGAGGACCTTACCCAGTCAATTTATTTATCCCAACTGCGG 35702  
QY 5404 TTAAGTGTGCTGAGTGGAGTCTTGGATACATCACTTTGAGTCAAACTCTGGATCTGC 5463  
DB |||||  
DB 35703 TTAAGTGTGCTGAGTGGAGTCTTGGATACATCACTTTGAGTCAAACTCTGGATCTGC 35762  
QY 5464 CAAAGGAACTTGAATAACAGGAGACCAAGCTAGCTATTTCTGTGAACCTCTAGAGGATT 5523  
DB |||||  
DB 35763 CAAAGGAACTTGAATAACAGGAGACCAAGCTAGCTATTTCTGTGAACCTCTAGAGGATT 35822  
QY 5524 TGGCCCTGCTCTTCAAAACAAACACAGGAGGAAAGTAACTAAAATCATAAATCCCCCATG 5583

DB 35823 TGGCCCTGCTCTTCAAAACAAACACAGGAGGAAAGTAACTAAAATCATAAAT -CCCCATG 35881  
QY 5584 GSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTVAACCTSTTTTCACTCTCATCTGCA 5643  
DB |||||  
DB 35882 GSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTTVAACCTSTTTTCACTCTCATCTGCA 35941  
QY 5644 CCCCCTCCATGCGCTGTATGACCAAGTACCTTACCCCTACCAAGAGTTTCTATGAGAAATG 5703  
DB |||||  
DB 35942 CCCCCTCCATGCGCTGTATGACCAAGTACCTTACCCCTACCAAGAGTTTCTATGAGAAATG 36001  
QY 5704 CAGCGTCCCGGAAATATTTGATGCGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCAACC 5763  
DB |||||  
DB 36002 CAGCGTCCCGGAAATATTTGATGCGCCCATCGTATAGGAGTCTTTTAAAGGAAACCCCAACC 36061  
QY 5764 TTCACTGCGCCACACCCCATATGCGCGCAACTGTCTATCACTCTGCCACTCTTTTGCATGAT 5823  
DB |||||  
DB 36062 TTCACTGCGCCACACCCCATATGCGCGCAACTGTCTATCACTCTGCCACTCTTTTGCATGAT 36121  
QY 5824 GCRAATCTCATTTATGACAGGAAATGAATTAATCCTAGTCTTCTGCGAGGACTTGA 5883  
DB |||||  
DB 36122 GCRAATCTCATTTATGACAGGAAATGAATTAATCCTAGTCTTCTGCGAGGACTTGA 36181  
QY 5884 GTCACTGTCTGTGCACTTACTTCCACCAAACTGGTATGTCTGATGGGGTGGAGTTCAA 5943  
DB |||||  
DB 36182 GTCACTGTCTGTGCACTTACTTCCACCAAACTGGTATGTCTGATGGGGTGGAGTTCAA 36241  
QY 5944 GATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCAACTCACCSGGGTATATGCG 6003  
DB |||||  
DB 36242 GATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCAACTCACCSGGGTATATGCG 36301  
QY 6004 ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCTCGGTACCCAT 6063  
DB |||||  
DB 36302 ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCTCGGTACCCAT 36361  
QY 6064 ACTCGCTGTGTAAAGCTTATTTTAAACCACTCTTGGGCTCCATGAGGTCTCGGCCCAA 6123  
DB |||||  
DB 36362 ACTCGCTGTGTAAAGCTTATTTTAAACCACTCTTGGGCTCCATGAGGTCTCGGCCCAA 36421  
QY 6124 AACCTCTAACTGTGTGATATGCTCCCTGTAACCTTTCARGGCAATATGTTTCAATCCCT 6183  
DB |||||  
DB 36422 AACCTCTAACTGTGTGATATGCTCCCTGTAACCTTTCARGGCAATATGTTTCAATCCCT 36481  
QY 6184 GTACTGAACTGGAACCACTTCCAGCACAGAAATTAACACCACTTCCGTTTCTAGTAGGA 6243  
DB |||||  
DB 36482 GTACTGAACTGGAACCACTTCCAGCACAGAAATTAACACCACTTCCGTTTCTAGTAGGA 36541  
QY 6244 CCTCTGTTTCCAAATSTGAAATTAACCCATACCTTCAAACTCACCTGTGTAAAAATTTAGC 6303  
DB |||||  
DB 36542 CCTCTGTTTCCAAATSTGAAATTAACCCATACCTTCAAACTCACCTGTGTAAAAATTTAGC 36601  
QY 6304 AATACTACATACAAACCACTCCCAATGCAATCAGGTGGGTAACCTCTCCCAACAAATA 6363  
DB |||||  
DB 36602 AATACTACATACAAACCACTCCCAATGCAATCAGGTGGGTAACCTCTCCCAACAAATA 36661  
QY 6364 GTCTGCTTACCCCTCAGGAAATTTTGTCTGTGTGCTTACCTCAGCTATCGTTTGTGAT 6423  
DB |||||  
DB 36662 GTCTGCTTACCCCTCAGGAAATTTTGTCTGTGTGCTTACCTCAGCTATCGTTTGTGAT 36721  
QY 6424 GGCTCTTCAGAAATCTATGTGCTTCTCTCATTTCTTAGTGCCCTCCCTATGCTACACT 6483  
DB |||||  
DB 36722 GGCTCTTCAGAAATCTATGTGCTTCTCTCATTTCTTAGTGCCCTTATGACCACTACACT 36781  
QY 6484 GAAACAGATTTTATACAGTTATGTCTATCTAAAGCCCGCAACAAAGAGTACCCATCTCT 6543  
DB |||||  
DB 36782 GAAACAGATTTTATACAGTTATGTCTATCTAAAGCCCGCAACAAAGAGTACCCATCTCT 36841  
QY 6544 CCTTTTGTATAGGACAGGAGTGTCTAGTGTCACTAGGTATCTGSCATTTGGCGGTATCACA 6603  
DB |||||  
DB 36842 CCTTTTGTATAGGACAGGAGTGTCTAGTGTCACTAGGTATCTGSCATTTGGCGGTATCACA 36901  
QY 6604 ACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTATAATGGGGAACATGGAACCGGCTC 6663

Db 36902 ACCTCTACTCAGTTCTACTCAAACTATCTCAAGAACTAAATGGGACATGACGGGTC 36961  
Qy 6664 GCCAGCTCCCTGGTCACTTGCAGATCAACTTAATCTCCCTAGCAGAGTAGTCTTCRA 6723  
Db 36962 GCCAGCTCCCTGGTCACTTGCAGATCAACTTAATCTCCCTAGCAGAGTAGTCTTCRA 37021  
Qy 6724 AATCGAAGAGCTTTAGACTTGTAAACCGCTGAGAGAGGGGAACTCTTTATTTTAAAGG 6783  
Db 37022 AATCGAAGAGCTTTAGACTTGTAAACCGCTGAGAGAGGGGAACTCTTTATTTTAAAGG 37081  
Qy 6784 GAGAAATGCTGTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 6843  
Db 37082 GAGAAATGCTGTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 37141  
Qy 6844 CSAGATCGAATACAACTAGACAGAGAGCTTGAACACTGACCTGGGCGCTCCTC 6903  
Db 37142 CGAGATCGAATACAACTAGACAGAGAGCTTGAACACTGACCTGGGCGCTCCTC 37201  
Qy 6904 AGCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6963  
Db 37202 AGCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 37261  
Qy 6964 CTCTCTTTGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7023  
Db 37262 CTCTCTTTGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 37321  
Qy 7024 GCTGTRAACTACAAATGGAGCCCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7083  
Db 37322 GCTGTRAACTACAAATGGAGCCCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 37381  
Qy 7084 CTGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7143  
Db 37382 CTGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 37441  
Qy 7144 GAAATCTCAGCTGCACAACTCTACTACGCCCCAAATTCAGCAGAGAGCAGTGTAGAGCGGT 7203  
Db 37442 GAAATCTCAGCTGCACAACTCTACTACGCCCCAAATTCAGCAGAGAGCAGTGTAGAGCGGT 37501  
Qy 7204 SGTGCGCAACCTCCCAACAGACACTGATGTTTAAATGACATCAAGGCACTGCTGCTGCTGCT 7263  
Db 37502 SGTGCGCAACCTCCCAACAGACACTGATGTTTAAATGACATCAAGGCACTGCTGCTGCTGCT 37561  
Qy 7264 AGACCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7323  
Db 37562 AGACCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 37621  
Qy 7324 ACATCCACCTTTAAACACCGGGCTTGCACCTTAAATGACATCAAGGCACTGCTGCTGCTGCTGCT 7383  
Db 37622 ACATCCACCTTTAAACACCGGGCTTGCACCTTAAATGACATCAAGGCACTGCTGCTGCTGCTGCT 37681  
Qy 7384 ACTAAATGCTAATTAAGGCAAGAGAGAGGAGTAAAGAAATAGCCAAATCAATCAATCAATCAAT 7443  
Db 37682 ACTAAATGCTAATTAAGGCAAGAGAGAGGAGTAAAGAAATAGCCAAATCAATCAATCAATCAAT 37741  
Qy 7444 AGAGCAGACAGAGAGGAGCAATGATCGGATATAAACCACAGTTCGAGCGCGGCAACGG 7503  
Db 37742 AGAGCAGACAGAGAGGAGCAATGATCGGATATAAACCACAGTTCGAGCGCGGCAACGG 37801  
Qy 7504 CAACCCCTTTGGGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7563  
Db 37802 CAACCCCTTTGGGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 37861  
Qy 7564 ATTAATCTTGCACCTGCT 7582  
Db 37862 ATTAATCTTGCACCTGCT 37880

RESULT 8  
HSAC000064  
LOCUS  
DEFINITION Human BAC clone RG083M05 from 7q21-7q22, complete sequence.  
AC000064  
AC000064.1 GI:1669369

KEYWORDS  
SOURCE  
ORGANISM

Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo;  
1 (bases 1 to 56093)  
Pauley, A.  
The sequence of H. sapiens BAC clone RG083M05  
Unpublished (1996)  
2 (bases 1 to 56093)  
Waterston, R.  
Direct Submission  
Submitted (13-NOV-1996)  
Genome Sequencing Center  
Department of Genetics, Washington University  
St. Louis, MO 63108, USA  
e-mail: sapiens@watson.wustl.edu

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

SOURCE INFORMATION:

This clone is from the first release of the human BAC library. The library contains cloned DNA from a human male fibroblast cell line 978SK. For references see: Shizuya et al., Proc. Natl. Acad. Sci. 89:8794-8797 (1992); Kim et al., Genomics 34:213-218 (1996).

VECTOR: pBelo

Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:

The orientation of this clone is unknown. Actual start of this clone is at base position 1 of H\_RG083M05; actual end is at 56093 of H\_RG083M05

This clone contains STS sws1725.

FEATURES  
Source

1..56093  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosomes="7"  
/maps="7q21-7q22"  
/clone="H RG083M05"  
/clone\_lib="CITB-978SK-B"  
complement(838..1131)  
/rpt\_family="ALU"  
-1360..16971  
/gene="WUGSC:H RG083M05.1"

repeat\_region

gene

CDS

join(1360..1503,4181..4370,4587..4774,6422..6556,9483..9547,11631..11773,11864..12021,13131..13296,14885..14988,16349..16546,16837..16971)  
/gene="WUGSC:H RG083M05.1"  
/note="ATPase; strong similarity to paroxisome biosynthesis protein PAB1 (P1D:51172019); coded for by human cDNA C04279 (NID:61467530)"  
/codon\_start=1  
/protein\_id="PAB46346.1"  
/db\_xref="GI:1669371"

/translation="KRLNIQKLEVAPEAWMOPSVLLDLDLILAGLPVPEHEH  
SPDAQRCEILCNVINKLDCDINKFTDLQHVAKETGGFVARDFTVLVRAHSLR  
SRQISITREKLVLITLDFQKALRGFLPASLRSLNHLKPRDLGWDKIGLHVRQILMD  
TILQPAKVLCKEYKPEFANLPITORTGILLYGPGTKTLGAGVIAESRMNFIS  
KGPILLSKYIGASEQAVRDIETRAQAAPCILLFDEFESIAPRGHNDTGTVDVVNQ



```
LLTLDGVEGLQGVVLAATSRPDLIDPALLRPGRLDKVCYPPPDVOTISYLSKTKO
QMLHSFLVSRLEINLVSDSLPLADDVLOHVASVTSFTGADLKALLYNAQLRALHG
MLKXMSILPDEKFNRYRYPFGSSYESELNGTSSDLSCQICLASPSMTQDLPGVP
KKQQFSPQPPVLTASQGBQELTQEQDQRADISIIKGYRSQSGEDEMNPQPGPI
KTRLAISHLMTALGHTRPRISIEDDKWKNFAEL"
/ complement (4948). .5130)
/rpt_family="ALU"
/ complement (6581). .7133)
/rpt_family="L1"
/ complement (7767). .8037)
/rpt_family="ALU"
/ complement (8186). .8472)
/rpt_family="ALU"
8473..8625
/ gene="WUGSC:H.RG083M05.1"
/ note="match to human 3', EST H75782 (NID:g1049794), bases
287-444"
8841..9161
/ gene="WUGSC:H.RG083M05.1"
/ note="match to human 5', EST H75921 (NID:g1050050), bases
21-348"
9481..9547
/ gene="WUGSC:H.RG083M05.1"
/ note="match to human 5', EST N22627 (NID:g1130501), bases
276-343"
/ complement (12612). .12907)
/rpt_family="ALU"
13670..13793
/ gene="WUGSC:H.RG083M05.1"
/ note="match to human 5', EST H41382 (NID:g17434), bases
143-266"
13794..13877
/rpt_family="ALU"
13878..13906
/ gene="WUGSC:H.RG083M05.1"
/ note="match to human 5', EST H41382 (NID:g917434), bases
20-58"
13907..14104
/rpt_family="ALU"
/ complement (14110). .14137)
/rpt_family="L1"
/ complement (15618). .15907)
/rpt_family="ALU"
17227..17522
/rpt_family="ALU"
18667..19235
/ note="match to human fetal brain 5', EST D61494
(NID:g970409), bases 1-255, and to human 3', EST R07476
(NID:g759393)"
19550..19670
/rpt_family="ALU"
21507..37303
/ note="similarity to various SS-RNA virus polyproteins;
pseudogene; region of matches and close matches to
multiple human ESTs, see R68740 (NID:g842257)"
37316..37489
/ inference="non-experimental evidence, no additional
details recorded"
/ note="Grail prediction, score = 80"
/ complement (38938). .39224)
/rpt_family="ALU"
39225..39707
/ note="match to multiple human ESTs, see N30113
(NID:g1148633)"
39800..40085
/rpt_family="ALU"
/ complement (40247). .40538)
/rpt_family="ALU"
/ complement (40632). .40924)
/rpt_family="ALU"
/ complement (42283). .42891)
/rpt_family="ALU"
/ complement (45474). .45613)

/rpt_family="ALU"
/ complement (45614). .45737)
/ note="match to human 3', EST H48898 (NID:g988738), bases
129-333"
/ complement (46107). .47026)
/ note="match to multiple human ESTs, see N81064
(NID:g1243765), H48897 (NID:g988737), and M78831
(NID:g273146)"
/ complement (47027). .47318)
/rpt_family="ALU"
/ complement (47365). .47782)
/ note="match to multiple human ESTs, see W37495
(NID:g1319089)"
47898..48115
/ note="match to human 5', EST H62306 (NID:g1015138), bases
93-368"
/ complement (48116). .48405)
/rpt_family="ALU"
/ complement (48406). .48584)
/ note="match to human 3', EST N29952 (NID:g1148472), bases
290-455, and 5', EST R12730 (NID:g765806)"
/ complement (48787). .49405)
/rpt_family="ALU"
/ complement (49406). .49534)
/ note="match to human 3', EST R65794 (NID:g838432), bases
309-440"
/ complement (49638). .49672)
/rpt_family="ALU"
/ complement (49674). .49890)
/ note="match to human 3', EST N29952 (NID:g1148472) and 5',
EST N29938 (NID:g1148458), sequences are from opposite
ends of the same clone"
/ complement (49698). .51806)
/ gene="WUGSC:H.RG083M05.2"
/ complement (join (49698). .49888, 51575). .51806))
/ gene="WUGSC:H.RG083M05.2"
/ note="coded for by human cDNAs W37389 (NID:g1319205),
R65891 (NID:g838529), R65794 (NID:g838432) and R65794
(NID:g838432)"
/ codon_start=1
/ protein_id="AAB46345.1"
/ db_xref="GI:1669370"
/ translation="MFVYFYFCGIIIFCPQGVYQIGDVSVIDEQDGKPYAQIRGF
IQDYCEKSAALTWLPTLSRQDFDPASVYIGPEDLPRKMEYLFVFCAPSSEYFK
SRSPFPVTPRPEKGIWTHVGTPTATIKESVANHL"
/ complement (51576). .51758)
/ gene="WUGSC:H.RG083M05.2"
/ inference="non-experimental evidence, no additional
details recorded"
/ note="Grail prediction, score = 86"

Query Match 86.1%; Score 6448; DB 5; Length 56093;
Best Local Similarity 95.7%; Pred. NO. 0;
Matches 6681; Conservative 195; Mismatches 50; Indels 53; Gaps 27;

QY 606 GATGGGAACCTTCCCGCAACAGCAAAACGCCCTTAAGACGTATTCTGGAATTTGGGA 665
|||||
Db 30953 GATGGGAACGTTCCTCCCAACAGCAAAACGCCCTTAAGACGTATTCTGGAATTTGGGA 31012
|||||
QY 666 MCAATTTGACCTCAGACACTAAGAAAGAACGACTTATTTCTTCTGCAGTCCGCCTG 725
|||||
Db 31013 CCAATTTGACCTCAGACACTAAGAAAGAACGACTTATTTCTTCTGCAGTCCGCCTG 31072
|||||
QY 726 GCATCTCTGAGGAAGTATAAATTATAACCATCTTACAGCTAGACATCTTTTGTAGAA 785
|||||
Db 31073 GCATCTCTGAGGAAGTATAAATTATAACCATCTTACAGCTAGACCTTTTGTAGAA 31132
|||||
QY 786 RAGGCAATGGAGTGAAGTGCATAGTACAACTTTCTTTTCAATTAAGAGACAATCTAC 845
|||||
Db 31133 RAGGCAATGGAGTGAAGTGCATAGTACAACTTTCTTTTCAATTAAGAGACAATCTAC 31192
|||||
QY 846 AATTATGTAAGAGTGTGATTATGCTTATGCTTACAGGAGCGCTTCAGAGTCTTACCTATC 905
|||||
```



Db	31193	AATTAATGTAAGAGTGATTTATGCTCCCTACAGAGCCCTTCAGAGTCTACCTCCCTATC	31252
Qy	906	CGAGCATCCCGAGTCTCTCCCMYTAATAAGAACCCCTTCACACCCAAATGGTCCAA	965
Db	31253	CCAGCATCCCGAGTCTCTCCCACTAAATAAGAGACCCCTTCACACCCAAATGGTCCAA	31312
Qy	966	AAGGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTCCCAATATTCCTCCCAATATGA	1025
Db	31313	AAGGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTCCCAATATTCCTCCCAATATGA	31372
Qy	1026	CCCTCCCAAGCAGTGGAGGAAAGAGAAATTCGGCCCAAGCCAGAGTGCATGTGCTTTTYY	1085
Db	31373	CCCTC-CCAAGCAGTGGAGGAAAGAGAAATTCGGCCCAAGCCAGAGTGCATGTGCTTTTTC	31431
Qy	1086	TCTCCAGACTTAAAGCAATAAACAAGACTTAGGTAAATTCAGATAAATTCAGATTAATG	1145
Db	31432	TCTCCAGACTTAAAGCAATAAACAAGACTTAGGTAAATTCAGATAAATTCAGATTAATG	31491
Qy	1146	CTATATTGRTTTTAAAGGGTTAGGACAAATTTTGTGATCTGACATGGAGAGATATA	1205
Db	31492	CTATATTGRTTTTAAAGGGTTAGGACAAATTTTGTGATCTGACATGGAGAGATATA-A	31550
Qy	1206	TGTCACCTGCTAAATCAGACACTAACCCCAATAGAGAGAGTGCACATAACTGCGAGCCT	1265
Db	31551	TGTCACCTGCTAAATCAGACACTAACCCCAATAGAGAGAGTGCACATAACTGCGAGCCT	31610
Qy	1266	GAGGTTTGGCGATCTCTGATATCTCAGTCAGTCAATGGATANGATGACAAAGAGG	1325
Db	31611	GAGGTTTGGCGATCTCTGATATCTCAGTCAGTCAATG-ATAGGATGACAAAG-AGG	31667
Qy	1326	AAAGAAATGATTTCCCAAGCCAGGCGAGCAGTTCCTCAGTCTASACCTCATTTGGGAC	1385
Db	31668	AAAGAAATGATTTCCCAAGCCAGGCGAGCAGTTCCTCAGTCTASACCTCATTTGGG---	31724
Qy	1386	ACAGAAATCAGTAAATGGAGATTTGCTGCTGAGACATTTGCTAACTTTGCTGTASAA	1445
Db	31725	ACAGAAATCAGACAT-GGAGATTTGCTGCTGAGACATTTGCTAACTTTGCTGTAGAA	31783
Qy	1446	GGACTAAGGAAACTAGGAAATCTAYGAAATCTCAATGATGTCCACCAATACACA	1505
Db	31784	GGACTAAGGAAACTAGGAAATCTCAATGATGTCCACCAATACACA	31842
Qy	1506	GGGAGGGAAGAAATCTTCTGCTGCTGAGAGACTAAGGAGGCAATGAGGAAGC	1565
Db	31843	GGGAGGGAAGAAATCTTCTGCTGCTGAGAGACTAAGGAGGCAATGAGGAAGC	31901
Qy	1566	GTGCTCTCTGTCACCTGCTCTTCTGAAGGCCAACTAATCTTAAAGCGTAACTTATCA	1625
Db	31902	GTGCTCTCTGTCACCTGCTCTTCTGAAGGCCAACTAATCTTAAAGCGTAACTTATCA	31961
Qy	1626	CTCAGTCAGTCGAGACATTTAG-AAAAAACTTCAAAAGTCTGCGTAGGCCCGGAGCAAA	1684
Db	31962	CTCAGTCAGTCGAGACATTTAGAAATAAACTTCAAAAGTCTGCGTAGGCCCGGAGCAAA	32021
Qy	1685	ACTTAGAAACCTTATTTGAACTTGGGAACTTGGGAACTTATATAGAGATCAGGAGAGC	1744
Db	32022	ACTTAGAAACCTTATTTGAACTTGGGAACTTGGGAACTTATATAGAGATCAGGAGAGC	32081
Qy	1745	AGGCGGACAGGACAAACGGGATTTAAATAAGGCCACCGCTTTAGTCAATGACCTCAGG	1804
Db	32082	AGGCGGACAGGACAAACGGGATTTAAATAAGGCCACCGCTTTAGTCAATGACCTCAGG	32141
Qy	1805	CAAGTGGACTTTGGAGGCTCTGGAAAGGAAAGCTGGGCAATTTGAATGCTTAATAGG	1864
Db	32142	CAAGTGGACTTTGGAGGCTCTGGAAAGGAAAGCTGGGCAATTTGAATGCTTAATAGG	32201
Qy	1865	GCTTGTCTCAGTGGGCTCTACAGGACACTTTTAAATAAGATTTGTCCAAAGTAAAGTAA	1924
Db	32202	GCTTGTCTCAGTGGGCTCTACAGGACACTTTTAAATAAGATTTGTCCAAAGTAAAGTAA	32261
Qy	1925	CGGCCCCCTGCTCATGCCCCCTTATTTCAAGGAAATCACTGGAAGGCCCACTGCCCGAGG	1984
Db	32262	CGGCCCCCTGCTCATGCCCCCTTATTTCAAGGAAATCACTGGAAGGCCCACTGCCCGAGG	32321
Qy	1985	GGACAAAGGCTTTTTCAGTCCAGAAAGCCACTAACAGATGATCCAGCAGCAGGACTGAGGG	2044
Db	32322	GGACAAAGGCTTTTTCAGTCCAGAAAGCCACTAACAGATGATCCAGCAGCAGGACTGAGGG	32381
Qy	2045	TGCTCTGGGCAAGCGCCATCCATGCTACCTCACAGAGCCCTGGGTATGCTTGACC	2104
Db	32382	TGCTCTGGGCAAGCGCCATCCATGCTACCTCACAGAGCCCTGGGTATGCTTGACC	32441
Qy	2105	ATTGAGGGCCAGGAAAGTTTGTCTCTGGACACTGGTGGGCTTTCTTAGTCTTACTCTTC	2164
Db	32442	ATTGAGGGCCAGG-AGGTTGTCTCTGGACACTGGTGGGCTTTCTTAGTCTTACTCTTC	32500
Qy	2165	TGTCCTGGGCAACACTGCTCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGACGGGCA	2224
Db	32501	TGTCCTGGGCAACACTGCTCTCCAGATCTGTCACTA-TCTGAGGGGGTCC-TAAGACGGGCA	32558
Qy	2225	GTCACTAGATACCTTTTCCAGGCCACTAAGTTATGACTGGGGAGCTTTTATCTTTTCAC	2284
Db	32559	GTCACTAGATAC-TTCTCCAGGCCACTAAGTTATG-ACTGGGGAGCTTTTATCTTTTCAC	32616
Qy	2285	ATGCTTTTCTAATTATGCTTGAAGCCCACTACCTTGTAGGGAGAGACATTTCTAGCAA	2344
Db	32617	ATGCTTTTCTAATTATGCTTGAAGCCCACTACCTTGTAGGGAGAGACATTTCTAGCAA	32676
Qy	2345	AAGCAGGGGCCATTTATACCTTGAAATAGGAGAGAAACACCCGTTTGTGTNCCCCTG	2404
Db	32677	AAGCAGGGGCCATTTATACCTTGAAATAGGAGAGAAACACCCGTTTGTGT-CCCCCTG	32735
Qy	2405	CTTGAGGAAGGATTAATCTGAACTCTGGGCAAGAGAGCAATATGACAGGCCAAA	2464
Db	32736	CTTGAGGAAGGATTAATCTGAACTCTGGGCAAGAGAGCAATATGACAGCG-CAAA	32794
Qy	2465	GAAATGCCCTCTGTTTCAAGTTAAACTAAAGGATTCACCTCTCTTTCCCTACCAAGGCA	2524
Db	32795	GAAATGCCCTCTGTTTCAAGTTAAACTAAAGGATTCACCTCTCTTTCCCTACCAAGGCA	32854
Qy	2525	GTACCCCTCTCAGACCCCAAGGCCCAACAGGATTTCCAAAGATTTGTTAAGGACTTAAAGC	2584
Db	32855	GTACCCCTCTCAGACCCCAAGGCCCAACAGGACTTCCAAAGATTTGTTAAGGACTTAAAGC	32914
Qy	2585	CCAAGCTTAGTAAACCAATGATTAACCTCCCTGAGTAATTTCCGTAGTGAATGAGAGG	2644
Db	32915	CCAAGCTTAGTAAACCAATGATTAACCTCCCTGAGTAATTTCCGTAGTGAATGAGAGG	32967
Qy	2645	CACAGAAACCCAGTCGAGAGTGGAGGTTAGTGCAGATCTCAGGATTAATCAATGAGGC	2704
Db	32968	CACAGAAACCCAGTCGAGAGTGGAGGTTAGTGCAGATCTCAGGATTAATCAATGAGGC	33025
Qy	2705	CGTTGCTCTTTATACCCAGCTGTACCTAGCCCTTATATCTGTGMYTTTCCCAATACAGA	2764
Db	33026	CGTTGCTCTTTATACCCAGCTGTACCTAGCCCTTATATCTGTGMYTTTCCCAATACAGA	33085
Qy	2765	GGAGCAGAGTGGTTTACASTCCTCGACCTTMAAGATGCTTCTTCTGATCCCTGTACA	2824
Db	33086	GGAGCAGAGTGGTTTACASTCCTCGACCTTMAAGATGCTTCTTCTGATCCCTGTACA	33145
Qy	2825	TCTGACTCTCAATTTCTGTTTGCCTTTGAAAGATCTTCAAAACCCCAATCTCAACTCAC	2884
Db	33146	TCTGACTCTCAATTTCTGTTTGCCTTTGAAAGATCTTCAAAACCCCAATCTCAACTCAC	33205
Qy	2885	CTGGACTTTTATACCCAGGGTTAGGGATAGYCCCACTTATTTTGGCCAGGCATTAGC	2944
Db	33206	CTGGACTTTTATACCCAGGGTTAGGGATAGYCCCACTTATTTTGGCCAGGCATTAGC	33265
Qy	2945	CCAAGACTTGAGCAATCTCTATACCTGGACA-CTTGTCTTCTGTGTAGTGGATGATTT	3004
Db	33266	CCAAGACTTGAGCAATCTCTATACCTGGACA-CTTGTCTTCTGTGTAGTGGATGATTT	33323
Qy	3005	ACTTTTGGCGCCCTTACAGAAACCTTGTGCCATCAAGCCACCCCAAGCCTCTTTMAATTT	3064
Db	33324	ACTTTTGGCGCCCTTACAGAAACCTTGTGCCATCAAGCCACCCCAAGCCTCTTTCAATTT	33383



Db	35523	CCTTAGCCAAATATCAAAAGTTCTTAAACATTTACAAGGAACCTATCCCTGAGAGAGG	35582
Qy	5284	GAAGAAGCAATTTCCACCCWGTGACATAGTATTAGTCAAGTCCTTCYCTCTAAATTTCCC	5343
Db	35583	GAAGAAGCAATTTCCACCCTTGTGACATGGTATTAGTCAAGTCCTTCCTCTAAATTTCCC	35642
Qy	5344	CATCCCTAGATACATCTCGGAGAGGCCCTTACCCAGTCAATTTATYATACCCCAACTGCGG	5403
Db	35643	CATCCCTAGATACATCTCGGAGAGGCCCTTACCCAGTCAATTTATYATACCCCAACTGCGG	35702
Qy	5404	TTAAAGTGGCTGGAGTGGATCTTTGGATACATCACACTTGGAGTCAAAATCCTGGATCTGC	5463
Db	35703	TTAAAGTGGCTGGAGTGGATCTTTGGATACATCACACTTGGAGTCAAAATCCTGGATCTGC	35762
Qy	5464	CAAGAAGCACTGAATTCAGGAGAGCAACGCTAGCTATTCCTGTGAACCTCTAGAGATT	5523
Db	35763	CAAGAAGCACTGAATTCAGGAGAGCAACGCTAGCTATTCCTGTGAACCTCTAGAGATT	35822
Qy	5524	TGCGCTGCTCTTCAAAACAAACAGGAGGAAGTAACATAATCATAAATCCCCCATG	5583
Db	35823	TGCGCTGCTCTTCAAAACAAACAGGAGGAAGTAACATAATCATAAATCCCCCATG	35881
Qy	5584	GSCTCCCTTATCATATTTTCTCTKTPASTGTTSTTTTACCTSTTTCACTCTCACTGCA	5643
Db	35882	GSCTCCCTTATCATATTTTCTCTKTPASTGTTSTTTTACCTSTTTCACTCTCACTGCA	35941
Qy	5644	CCCCCTCCATGCGCTGTATGACAGTAGTCCCTCCCTVACMAAGAGTTCTATAGGAATG	5703
Db	35942	CCCCCTCCATGCGCTGTATGACAGTAGTCCCTCCCTVACMAAGAGTTCTATAGGAATG	36001
Qy	5704	CAGCGTCCCGGAATATTTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCAAC	5763
Db	36002	CAGCGTCCCGGAATATTTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCAAC	36061
Qy	5764	TTCACTGCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTTGCATGCA	5823
Db	36062	TTCACTGCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTTGCATGCA	36121
Qy	5824	GCAAAATCACTATTTGGACAGAAATAATGATTAATCTCTAGTTGTCTGGAGACTTGA	5883
Db	36122	GCAAAATCACTATTTGGACAGAAATAATGATTAATCTCTAGTTGTCTGGAGACTTGA	36181
Qy	5884	GTCACTGTCTGTGGACTTACTTCAACCAACTGGTATGCTGATCGGGGTGAGTTCAA	5943
Db	36182	GTCACTGTCTGTGGACTTACTTCAACCAACTGGTATGCTGATCGGGGTGAGTTCAA	36241
Qy	5944	GATCAGGCAAGAGAAAAATGTAAAGAGTAATCTCCCAACTCAACCGGGTACATGGC	6003
Db	36242	GATCAGGCAAGAGAAAAATGTAAAGAGTAATCTCCCAACTCAACCGGGTACATGGC	36301
Qy	6004	ACCTTAGCCCTTACAAAGSACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCCAT	6063
Db	36302	ACCTTAGCCCTTACAAAGSACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCCAT	36361
Qy	6064	ACTCGCTGGTAAGCTATTTAATACACCTCACTGGGTCATGAGTCTCGGCCAA	6123
Db	36362	ACTCGCTGGTAAGCTATTTAATACACCTCACTGGGTCATGAGTCTCGGCCAA	36421
Qy	6124	AACCCCTACTAATGTGTGGATATGSCCTCCCTGAACTTCARGCCATATGTTTCAATCCCT	6183
Db	36422	AACCCCTACTAATGTGTGGATATGSCCTCCCTGAACTTCARGCCATATGTTTCAATCCCT	36481
Qy	6184	GTAACCTGAACAATGGAAACAATTTGACACAGAAATAAACCACCTTCCTGTTTGTAGGA	6243
Db	36482	GTAACCTGAACAATGGAAACAATTTGACACAGAAATAAACCACCTTCCTGTTTGTAGGA	36541
Qy	6244	CCTCTGTTTCCAAATSTGAAATAACCCATACCTCAAACTCACTGTGTAAATTTAGC	6303
Db	36542	CCTCTGTTTCCAAATSTGAAATAACCCATACCTCAAACTCACTGTGTAAATTTAGC	36601
Qy	6304	AATACTACATACACACCACTCCCAATGCATCAGGTGGGTAACTCTCCACACAAATA	6363
Db	36602	AATACTACATACACACCACTCCCAATGCATCAGGTGGGTAACTCTCCACACAAATA	36661
Qy	6364	GTCTGCTACCTCCCTCAGGAATAATTTTTTGTCTGTGGTACCTCAGCCCTATCGTTGTTGAAT	6423
Db	36662	GTCTGCTACCTCCCTCAGGAATAATTTTTTGTCTGTGGTACCTCAGCCCTATCGTTGTTGAAT	36721
Qy	6424	GGCTCTTTCAGAAATCTATGTCTCTCTCAATTTCTAGTGGCCCTCATGRCATCTACCT	6483
Db	36722	GGCTCTTTCAGAAATCTATGTCTCTCTCAATTTCTAGTGGCCCTCATGRCATCTACCT	36781
Qy	6484	GAACAAGATTTATACAGTTATGTATATCTAAAGCCCGCAACAAAGAGAGTACCATTTCT	6543
Db	36782	GAACAAGATTTATACAGTTATGTATATCTAAAGCCCGCAACAAAGAGAGTACCATTTCT	36841
Qy	6544	CTTTTGTATATAGGAGCAGAGTGCTAGTGCTAGTGTCTAGTGTCTGGCATTTGGCGGTATCACA	6603
Db	36842	CTTTTGTATATAGGAGCAGAGTGCTAGTGCTAGTGTCTAGTGTCTGGCATTTGGCGGTATCACA	36901
Qy	6604	ACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAAACGGGTC	6663
Db	36902	ACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAAACGGGTC	36961
Qy	6664	GCGACTCCCTGCTGCTACCTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTCRA	6723
Db	36962	GCGACTCCCTGCTGCTACCTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTCRA	37021
Qy	6724	AATCGAAGAGCTTTAGACTYGTCTAACCGCTGAPAGAGGGGAACTGTGTTTATTTTAGGG	6783
Db	37022	AATCGAAGAGCTTTAGACTYGTCTAACCGCTGAPAGAGGGGAACTGTGTTTATTTTAGGG	37081
Qy	6784	GAAGATGCTGTATTTATGTTTAACTCCGGAATCGTCACTGAGAAAGTTAAGAAATTT	6843
Db	37082	GAAGATGCTGTATTTATGTTTAACTCCGGAATCGTCACTGAGAAAGTTAAGAAATTT	37141
Qy	6844	CSAGATCGAATAACAAGTAKAGCAGARGAGCTTCGAAACACACTGGACCTCGGGCCCTCCTC	6903
Db	37142	CSAGATCGAATAACAAGTAKAGCAGARGAGCTTCGAAACACACTGGACCTCGGGCCCTCCTC	37201
Qy	6904	AGCCATGATGCCCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTA	6963
Db	37202	AGCCATGATGCCCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTA	37261
Qy	6964	CTCTCTTTGGACCTCTGATCTTTTACCTCTTTTAACTTTGTCTCTTCCAGAAATCGAA	7023
Db	37262	CTCTCTTTGGACCTCTGATCTTTTACCTCTTTTAACTTTGTCTCTTCCAGAAATCGAA	37321
Qy	7024	GCTGTAAACTACAAATGAGGCCCAAGATGCAAGTCAAGACTAAAGATCTACCGCAGACCC	7083
Db	37322	GCTGTAAACTACAAATGAGGCCCAAGATGCAAGTCAAGACTAAAGATCTACCGCAGACCC	37381
Qy	7084	CTGGACCGGCTGTAGCCACGATCTGATGTTTAAATGACATCAAAGGCAACCCCTCCTGAG	7143
Db	37382	CTGGACCGGCTGTAGCCACGATCTGATGTTTAAATGACATCAAAGGCAACCCCTCCTGAG	37441
Qy	7144	GAATCTCAGTGTGCAACCTCTAOCGCCCAATTCAGCAGGAGCAGTAGTAGAGCGGT	7203
Db	37442	GAATCTCAGTGTGCAACCTCTAOCGCCCAATTCAGCAGGAGCAGTAGTAGAGCGGT	37501
Qy	7204	SGTGGGCCAACTCCCAACACAGCTTAGGTTTCTCTGTTGATGGGGGACGTGAGAGAC	7263
Db	37502	SGTGGGCCAACTCCCAACACAGCTTAGGTTTCTCTGTTGATGGGGGACGTGAGAGAC	37561
Qy	7264	AGGACTAGCTGATTTCTTAGGCTGATTAAGAAATCCYTAAGCCTTAGSTGGGAAGGTGACC	7323
Db	37562	AGGACTAGCTGATTTCTTAGGCTGATTAAGAAATCCYTAAGCCTTAGSTGGGAAGGTGACC	37621
Qy	7324	ACATCACTTTTAAACAGCGGCTTTGCAACTTAGVTCAACCTGACCAATCAGAGAGCTC	7383
Db	37622	ACATCACTTTTAAACAGCGGCTTTGCAACTTAGVTCAACCTGACCAATCAGAGAGCTC	37681
Qy	7384	ACTAAATGCTTAATTTAGGCAAGACAGGAGGTAAAGAAATACCAATCATYTTATTTGCTG	7443
Db	37682	ACTAAATGCTTAATTTAGGCAAGACAGGAGGTAAAGAAATACCAATCATYTTATTTGCTG	37741











Db	87860	AAACCTATGCCACTCGAGGGGACCTTTTAGAGGTTCCTTTGACTGATCCCGACC-TCAAC	87802
Qy	3905	TTGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGAGTGG	3964
Db	87801	TTGTATACTGATGGAAGTTCCTTTGTAGAAAAGAGGACTTCGAAAAGTGGGGTATGAGTGG	87742
Qy	3965	GTCACTGATTAATGGAATATTTGAAAGTAATTCCTCTCATCCAGGAACCTAGTCTYAGCTTR	4024
Db	87741	GTCACTGATTAATGGAATATTTGAAAGTAATTCCTCTCATCCAGGAACCTAGTCTYAGCTTR	87682
Qy	4025	GCAGAACTAATAGCCYCAATYKGGGCACTAGAAATTAGGAGAGCAAAAAGGGYAAATATA	4084
Db	87681	GCAGAACTAATAGCCCTCACTTGGGCACTAGAAATTAGGAGAGCAAAAAGGGYAAATATA	87622
Qy	4085	TATACAGACTCTRARTATGCTYACCTAGTCTNTCCATGCCCATGMRCAATATGSRAGAA	4144
Db	87621	TATACAGACTCTAATATATGCTTACCTAGTCTCCTCATGCCCATGMRCAATATGSRAGAA	87562
Qy	4145	AGGGAATTCCTAATCTTCYAGRGAAACCTTATCAMAATCATCAGGAAGCCATTAGGAAATTA	4204
Db	87561	AGGGAATTCCTAATCTTCYAGRGAAACCTTATCAMAATCATCAGGAAGCCATTAGGAAATTA	87502
Qy	4205	TTATYTGCGWGTACAGAAACCTTARAGAGTGGAGTCTTACACTGCGGGGTCAATCANAAA	4264
Db	87501	TTATYTGCGWGTACAGAAACCTTARAGAGTGGAGTCTTACACTGCGGGGTCAATCANAAA	87442
Qy	4265	GGAAAGAAAGGGAAATASAGRGAAATGCAAGCAKATATTTGAAGCAAAAGAGCTGCA	4324
Db	87441	GGAAAGAAAGGGAAATASAGRGAAATGCAAGCAKATATTTGAAGCAAAAGAGCTGCA	87382
Qy	4325	AGGCAGACCTCTCATTTAGAAATGCTTTATTAACCTTCCCTTAGTATAGGGTAAATCCCTTC	4384
Db	87381	AGGCAGACCTCTCATTTAGAAATGCTTTATTAACCTTCCCTTAGTATAGGGTAAATCCCTTC	87322
Qy	4385	CGGGAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGAACTTCAAGG-CAG	4443
Db	87321	CGGGAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGAACTTCAAGGAGACAG	87262
Qy	4444	TTTTTCTCCCTCGGGACGGTTAGCCACTGAAGAGGGGAAATATCTTTTCCTGCAACTAT	4503
Db	87261	TTTTTCTCCCTCGGGACGGTTAGCCACTGAAGAGGGGAAATATCTTTTCCTGCAACTAT	87202
Qy	4504	CCAATGGAATTACTTAAACCCCTTATCAAAACCTTTTCACTTAGGATCGATGACCA	4563
Db	87201	CCAATGGAATTACTTAAACCCCTTATCAAAACCTTTTCACTTAGGATCGATGACCA	87142
Qy	4564	TCARATGGCAATCATTTACTGCAAGAGCCCTTTTCAAAACCTATCAAGGARATAT	4623
Db	87141	TCARATGGCAATCATTTACTGCAAGAGCCCTTTTCAAAACCTATCAAGGARATAT	87082
Qy	4624	CAGGGCTGTGAAGTGTCCARAAATATCCCTGCTCTATCGGCAAGCTTCTTCAGG	4683
Db	87081	CAGGGCTGTGAAGTGTCCARAAATATCCCTGCTCTATCGGCAAGCTTCTTCAGG	87022
Qy	4684	ARAACAAARACAGGCCATTAACCTGARAARACTGGCAACTGATTTTACCAAGCC	4743
Db	87021	ARAACAAARACAGGCCATTAACCTGARAARACTGGCAACTGATTTTACCAAGCC	86962
Qy	4744	AAACCTCAGGATTTTCACTATCTACTAGTCTGGGTATATCTTTCAAGGGTGGGAGAG	4803
Db	86961	AAACCTCAGGATTTTCACTATCTACTAGTCTGGGTATATCTTTCAAGGGTGGGAGAG	86902
Qy	4804	GCCTTCCCTGTAGGACAGAAAAAGGCCCAAGAGGTAAATAAGGCACTAGTTCAATGAATA	4863
Db	86901	GCCTTCCCTGTAGGACAGAAAAAGGCCCAAGAGGTAAATAAGGCACTAGTTCAATGAATA	86842
Qy	4864	ATTCCCAATTCGGACTTTCCTCCGAGGCTTACAGAGTGACAATAGCCCTCTTTCAGGCC	4923
Db	86841	ATTCCCAATTCGGACTTTCCTCCGAGGCTTACAGAGTGACAATAGCCCTCTTTCAGGCC	86782
Qy	4924	ACAGTAACCCAGGGAGTATCCCAAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGA	4983
Db	86781	ACAGTAACCCAGGGAGTATCCCAAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGA	86722

QY 6064 ACTCGCTGGTAAGCTATTTAATACCAACCTCACTGGGCTCCATGAGGTCTCGGCCAA 6123  
DB 85642 ACTCGCTGGTAAGCTATTTAATACCAACCTCACTGGGCTCCATGAGGTCTCGGCCAA 85583  
QY 6124 AACCTACTAACTGTGGATATGCTCCCTGAACTTCARGCCATATGTTTCAATCCCT 6183  
DB 85582 AACCTACTAACTGTGGATATGCTCCCTGAACTTCARGCCATATGTTTCAATCCCT 85523  
QY 6184 GTACCTGAACTGAACTTTCAGCACAGAAATTAACACCACTTCGGTTTCTAGTGA 6243  
DB 85522 GTACCTGAACTGAACTTTCAGCACAGAAATTAACACCACTTCGGTTTCTAGTGA 85463  
QY 6244 CCTCTGTTTCCAACTGGAATAAACCCATACCTCAACCTCACCTGTGTAAATTTAGC 6303  
DB 85462 CCTCTGTTTCCAACTGGAATAAACCCATACCTCAACCTCACCTGTGTAAATTTAGC 85403  
QY 6304 AATACTACATACACAACTCCCAATGATCAGTGGGTAACTCCTCCCAACAAATA 6363  
DB 85402 AATACTACATACACAACTCCCAATGATCAGTGGGTAACTCCTCCCAACAAATA 85343  
QY 6364 GTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCTATGCTGTTTGAAT 6423  
DB 85342 GTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCTATGCTGTTTGAAT 85283  
QY 6424 GCTCTTTCAGAACTATGCTCTCTCATTTCTAGTGCCTCCCVATGRCATCTACACT 6483  
DB 85282 GCTCTTTCAGAACTATGCTCTCTCATTTCTAGTGCCTCCCVATGRCATCTACACT 85223  
QY 6484 GAACAAGATTATACAGTTATGCTCATATCTAAGCCCGCAACAAAGAGTACCCATTTCT 6543  
DB 85222 GAACAAGATTATACAGTTATGCTCATATCTAAGCCCGCAACAAAGAGTACCCATTTCT 85163  
QY 6544 CTTTTTGTATAGGAGCAGGAGTGTAGTGCACTAGGTACTGGCAATGGCGGTATACCA 6603  
DB 85162 CTTTTTGTATAGGAGCAGGAGTGTAGTGCACTAGGTACTGGCAATGGCGGTATACCA 85103  
QY 6604 ACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGT 6663  
DB 85102 ACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGT 85043  
QY 6664 GCCGACTCCCTGGTCAAGATCAACTTAATCTCCCTAGCAGCAGTACTCTTCRA 6723  
DB 85042 GCCGACTCCCTGGTCAAGATCAACTTAATCTCCCTAGCAGCAGTACTCTTCRA 84983  
QY 6724 AATCGAAGAGCTTTAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6783  
DB 84982 AATCGAAGAGCTTTAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 84923  
QY 6784 GAAGATGCTGTTATGTTATGTTATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 6843  
DB 84922 GAAGATGCTGTTATGTTATGTTATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 84863  
QY 6844 CSAGATCGAATAACAGTACAGAGAGCTTCGAAACACCTGGACCTCGGGGCTCCTC 6903  
DB 84862 CGAGATCGAATAACAGTACAGAGAGCTTCGAAACACCTGGACCTCGGGGCTCCTC 84803  
QY 6904 AGCCATAGGAGCCCTGGATTTCCCTTTCTAGGACCTCTAGCAGCTATATATGCTA 6963  
DB 84802 AGCCATAGGAGCCCTGGATTTCCCTTTCTAGGACCTCTAGCAGCTATATATGCTA 84743  
QY 6964 CTCCTCTTTGGACCTGTTATCTTTACCTCTGTTTAACTTTGCTCTTCAGAAATCGAA 7023  
DB 84742 CTCCTCTTTGGACCTGTTATCTTTAACTCTGTTTAACTTTGCTCTTCAGAAATCGAA 84683  
QY 7024 GCTGTAAACTACAAATGAGCCCAAGATGCACTCAAGCTAAAGTCTACCGCAGACCC 7083  
DB 84682 GCTGTAAACTACAAATGAGCCCAAGATGCACTCAAGCTAAAGTCTACCGCAGACCC 84623  
QY 7084 CTGGAACCGGCTGTTAGCCCAAGATGTTATGATGATCAATCAAGGACCCCTCTGAG 7143  
DB 84622 CTGGAACCGGCTGTTAGCCCAAGATGTTATGATGATCAATCAAGGACCCCTCTGAG 84563  
QY 7144 GAAATCTCAGTGCACAACTCTACTACGCCCAATTCAGAGGAGCAGTTAGACGGT 7203

DB 84562 GAAATCTCAGTGCACAACTCTACTACGCCCAATTCAGAGGAGCAGTTAGACGGT 84503  
QY 7204 SGTGGCCAACTCCCAACAGCACTTAGGTTTCTGTTGAGATGGGGGACTGAGAGAC 7263  
DB 84502 CGTGGCCAACTCCCAACAGCACTTAGGTTTCTGTTGAGATGGGGGACTGAGAGAC 84443  
QY 7264 AGGACTAGCTGATTTCTCTAGGCTGATTAAGAACTCCYTAAGCCTAGSTAGGAAAGCTGACC 7323  
DB 84442 AGGACTAGCTGATTTCTCTAGGCTGATTAAGAACTCCYTAAGCCTAGGAAAGCTGACC 84383  
QY 7324 ACATCCACTTTAAACACGGGCTTGCACCTTAGTTCACCTGACCTACCAATCAGAGAGCTC 7383  
DB 84382 ACATCCACTTTAAACACGGGCTTGCACCTTAGTTCACCTGACCTACCAATCAGAGAGCTC 84323  
QY 7384 ACTAAATGCTAATTAGCAAAAGACAGAGGTAAGAAATAGCCCAATCATYATTTGCMTG 7443  
DB 84322 ACTAAATGCTAATTAGCAAAAGACAGAGGTAAGAAATAGCCCAATCATCTATTGCTG 84263  
QY 7444 AGAGCAGCAGGAGGAGCAATGATCGGGATATAAACCAGTCTTCAGCGCGCAACGG 7503  
DB 84262 AGAGCAGCAGGAGGAGCAATGATCGGGATATAAACCAGTCTTCAGCGCGCAACGG 84203  
QY 7504 CAACCCCTTTGGGTCCTCCCTTTGATGCGGAGCTCTGTTTTCATGCTATTTCACTCT 7563  
DB 84202 CAACCCCTTTGGGTCCTCCCTTTGATGCGGAGCTCTGTTTTCATGCTATTTCACTCT 84143  
QY 7564 ATTAATCTTCARCTGCR 7582  
DB 84142 ATTAATCTTCGCACTGCA 84124

RESULT 10  
AY101583  
LOCUS  
DEFINITION Homo sapiens isolate 21 endogenous retrovirus HERV-W, ERVWE1 locus,  
allele B, complete sequence.  
ACCESSION AY101583  
VERSION AY101583.1 GI:37544401  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 10222)  
AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,  
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.  
The endogenous retroviral locus ERVWE1 is a bona fide gene involved  
in hominoid placental physiology  
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)  
14757826  
JOURNAL PUBMED  
PUBMED  
REFERENCE 2 (bases 1 to 10222)  
AUTHORS Mallet,F., Bouton,O. and Oriol,G.  
TITLE Direct Submission  
JOURNAL Submitted (06-MAY-2002) Retrovirology Department, UMR 2142  
CNRS-bioMérieux, Ecole Normale Supérieure de Lyon - 46 allée  
d'Italie, Lyon 69364 cedex 07, France  
FEATURES  
Location/Qualifiers  
1..10222  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/isolate="21"  
/isolation\_source="PBMC of individual 21"  
/db\_xref="taxon:9606"  
/chromosome="7"  
/map="7q21-q22"  
/sex="female"  
/cell\_type="PBMC"  
/note="isolated by PCR; ERVWE1 locus, allele B  
endogenous\_virus: HERV-W"  
1..780  
join(248..811,2887..9813)

LTR  
mRNA

```

/note="putative mRNA transcript 1; gag-pol pseudogene"
join(248..811,7585..9813)
/product="envelope glycoprotein"
/note="putative mRNA transcript 2"
join(248..811,9248..9813)
/note="putative mRNA transcript 3; no coding region"
7812..9428
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAM68162.1"
/db_xref="GI:37544402"
/tranlation="MALPYHIEFLFTVLLPSFTLTAPPCCMTSSPYQBFPLMRMORP
GNIDAYSRLSGTPTFTATHENPNCYHSAILCMHANTHTWTGKINSCPGGLGV
TVCWYFTQMGSDGGVQDOAREKHAKEVISQLTRVHGTSYKGLDLSKLEHTLRT
HTRLVSNLTTLTGLHSEVQNTNCWICLPLNFRPYSIPIVPEQWNNEFTLNTVS
LVGLPLNLTBITHTLSLTCVAKTSPTNVTQIPLNFRVYTPQVCLPSPGIPFVCVGS
YRLCNGSSMSFLTVPLVPMWTIVTEQDLSYVVISKPRNKRVPILPVTIGAGVLGALG
TLGIGTSTQFYFKLSQELNGDMERVADSLVTLQDLQSLAAVLQNRPALDLITAE
RGCTCLFLGCECCYYNQSGIVTEKYKEIRDRIQRAABELNRTGPNGLSLSQWMPWILIP
FLGLRAAILLULLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRRLPLDRPASP
RSDVNDIKGTPPBEISAAQPLLRPNPNSAGGS"
9487..10222

```

Query Match	86.1%	Score 6446.4	DB 5	Length 10222
Best Local Similarity	95.7%	Pred. No. 0		
Matches 6680	Conservative 195	Mismatches 51	Indels 53	Gaps 27
Qy	606	GATGGGAAACGTTCCCGCAGACAAAAACGCCCTTAAGACGTTATCTCGAATAATTGGGA	665	
Db	2886	GATGGGAAACGTTCCCGCAGACAAAAACGCCCTTAAGACGTTATCTCGAATAATTGGGA	2945	
Qy	666	MAATTTGACCCCTCAGACACTTAAGAAAGAAACGACTTATATCTTCCTGACGTGCGCGCTG	725	
Db	2946	CAATTTGACCCCTCAGACACTTAAGAAAGAAACGACTTATATCTTCCTGACGTGCGCGCTG	3005	
Qy	726	GCATCTCTGAGGGAAGTATAAATATATAACACCATTTTACAGCTAGACYTCTTTTGTAGAA	785	
Db	3006	GCATCTCTGAGGGAAGTATAAATATATAACACCATTTTACAGCTAGACCTCTTTTGTAGAA	3065	
Qy	786	AAGCAAAATCGAGTGAAGTCCCATNAGTACAAACTTCTTTTCAATTAAGAGACAATCTCAC	845	
Db	3066	AAGCAAAATCGAGTGAAGTCCCATNAGTACAAACTTCTTTTCAATTAAGAGACAATCTCAC	3125	
Qy	846	AAATTATGTAAAAAGTGTGATTTATGCCCTACAGAGAGCCCTTCAGAGTCTACCTCCCTATC	905	
Db	3126	AAATTATGTAAAAAGTGTGATTTATGCCCTACAGAGAGCCCTTCAGAGTCTACCTCCCTATC	3185	
Qy	906	CCAGCATCCCGACTCCTCTCCCAAMTAAATAAGAGACCCGCCCTTCAACCCAAATGGTCCAA	965	
Db	3186	CCAGCATCCCGACTCCTCTCCCAACTAATAAGAGACCCGCCCTTCAACCCAAATGGTCCAA	3245	
Qy	966	AAGGAGATGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCTAATTTCCCAATATGA	1025	
Db	3246	AAGGAGATGACAAAAGGGTAAACAGTGAACCAAAGAGTGCCTAATTTCCCAATATGA	3305	
Qy	1026	CCCTCCCAAGCAGTGGGAGAGAGAAATTCGGCCCAAGCAGAGTGCATGTGCTTTTTYY	1085	
Db	3306	CCCTC-CCAAGCAGTGGGAGAGAGAAATTCGGCCCAAGCAGAGTGCATGTGCTTTTTTC	3364	
Qy	1086	TCTCCACAGCTTAAAGCAAAATAAAAACAGACTTTAGGTAAATTTCTCAGATAAATCTGATGG	1145	
Db	3365	TCTCCACAGCTTAAAGCAAAATAAAAACAGACTTTAGGTAAATTTCTCAGATAAATCTGATGG	3424	
Qy	1146	CTATATTTGRTGTTTTACAAGGGTTAGGACAAATTTCTTTGATCTGACATGGAGAGATATA	1205	
Db	3425	CTATATTTGRTGTTTTACAAGGGTTAGGACAAATTTCTTTGATCTGACATGGAGAGATATA	3483	
Qy	1206	TGTCACGTCTAAATCAGACACTAACCCCAATAAGAGAGAGTGCACCATTAATCTGACGCT	1265	
Db	3484	TGTCACGTCTAAATCAGACACTAACCCCAATAAGAGAGAGTGCACCATTAATCTGACGCT	3543	

Qy	1266	GAGRGTTTGCGATCTCTGCTATCTCTCAGTCAGGTCAATGGATANGATGACAA	CAGAAGG	1326	
Db	3544	GAGAGTTTGCGATCTCTGCTATCTCTCAGTCAGGTCAATG	---ATAGGATGACAA	CAG-AGG	3600
Qy	1326	AAAGNAAATGATTTCCCCACAGGCCAGCARGCAGTTC	CCAGTCTTASACCCCTCATTTGGGAC	1385	
Db	3601	AAAGAGAAATGATTTCCCCACAGGCCAGCAGCAGTTC	CCAGTCTAGACCCCTCATTTGGG- --	3657	
Qy	1386	ACAGAAATCAGTAACATGGGAGATTGGTCTGCAGACATTT	GTCTAACTTGCTGTGCTASAA	1445	
Db	3658	ACACAGAAATCAGAAACAT- GGAGATTGGTCTGCAGACATTT	GTCTAACTTGCTGTGCTAGAA	3716	
Qy	1446	GGACTAAGGAAAACTATAGGAAAGAAARTCTATGAAATTA	CTCAATGATGTCCACCAATAACACA	1505	
Db	3717	GGACTAAGGAAAACTATAGGAAAGAAARTCTATGAAATTA	CTCAATGATGTCCACCAATAACACA	3775	
Qy	1506	GGGGAGGGGAAGAAATCTTACTGCTTTCTGGAGAGACTAA	AGGGAGGCATTCAGGAAGC	1565	
Db	3776	GGGAAGGGGAAGAAATCTTACTGCTTTCTGGAGAGACTAA	AGGGAGGCATTCAGGAAGC	3834	
Qy	1566	GTGCCTCTCTGTCAACCTGACTCTTCTGAAGGCCAACTAA	TCCTTAAAGCGTAACTTTATCA	1625	
Db	3835	GTGCCTCTCTGTCAACCTGACTCTTCTGAAGGCCAACTAA	TCCTTAAAGCGTAACTTTATCA	3894	
Qy	1626	CTCAGTCAGCTGCAGACATTTAG- AAAAATCTTCAAAGTCT	CGGTAGGCCCGGAGCAAA	1684	
Db	3895	CTCAGTCAGCTGCAGACATTTAGAAAAAACTTCAAAGTCT	CGGTAGGCCCGGAGCAAA	3954	
Qy	1685	ACTTAGAAACCTTATTTGAACCTTGGCACTTCGCTTTT	TATTAATAGAGATCAGGAGGAGC	1744	
Db	3955	ACTTAGAAACCTTATTTGAACCTTGGCAACCTTCGCTTTT	TATTAATAGAGATCAGGAGGAGC	4014	
Qy	1745	AGCGGAAACAGGACAAACGGGATTTAAAAAAGGCCACAC	CGCTTTTAGTCATGACCCCTCAGG	1804	
Db	4015	AGCGGAAACAGGACAAACGGGATTTAAAAAAGGCCACAC	CGCTTTTAGTCATGACCCCTCAGG	4074	
Qy	1805	CAAGTGGACTTTGGAGGCTCTGAAAAGGAAAGCTGGGCAA	AAATTTGAATTCCTAATAGG	1864	
Db	4075	CAAGTGGACTTTGGAGGCTCTGAAAAGGAAAGCTGGGCAA	AAATTTGAATTCCTAATAGG	4134	
Qy	1865	GCCTTGCTTCCAGTGGCGTCTACAAGACACTTTAAAAA	AGATTGTTCCAGTAGAAGTTAG	1924	
Db	4135	GCCTTGCTTCCAGTGGCGTCTACAAGACACTTTAAAAA	AGATTGTTCCAGTAGAAGTTAG	4194	
Qy	1925	CCGCCCTTCGTCCATGCCCTTATTTCAAGGAAATCACTT	GGAAGGCCACTGCCCCGAGG	1984	
Db	4195	CCGCCCTTCGTCCATGCCCTTATTTCAAGGAAATCACTT	GGAAGGCCACTGCCCCGAGG	4254	
Qy	1985	GGACAAAGGTCTTTTGAGTCAGAGGCACCTAACCCAGAT	GATCCAGCAGCAGGACTGAGGG	2044	
Db	4255	GGACAAAGGTCTTTTGAGTCAGAGGCACCTAACCCAGAT	GATCCAGCAGCAGGACTGAGGG	4314	
Qy	2045	TGCCTGGGGCAAGGCCATCCATGCCATCACCTCACAGAG	CCCTGGGTATGCTTTGACC	2104	
Db	4315	TGCCTGGGGCAAGGCCATCCATGCCATCACCTCACAGAG	CCCTGGGTATGCTTTGACC	4374	
Qy	2105	ATTGAGGGCCAGGAAGTGTCTCTCGGACACTGGTGGCT	CTCTTTAGTCTTACTCTTTC	2164	
Db	4375	ATTGAGGGCCAGG- AGTGTGTCTCTCGGACACTGGTGG	CTCTCTTTAGTCTTACTCTTTC	4433	
Qy	2165	TGTCCCGGACAACTGTCTCCAGATCTGTCACTATTCTG	AGGGGGTCCNTAAGACGGGCA	2224	
Db	4434	TGTCCCGGACAACTGTCTCCAGATCTGTCACTA- TCTG	AGGGGGTCC- TAAGACGGGCA	4491	
Qy	2225	GTCACTAGATACATTTTTCAGGCCACTAAGTTATGAACT	TGGGAGCTTTTATCTTTTCAC	2284	
Db	4492	GTCACTAGATAC- TTCTCCAGCCACTAAGTTATG- ACT	TGGGAGCTTTTATCTTTTCAC	4549	
Qy	2285	ATGCTTTTCTAAATTATGCTTGAAGGCCCACTACCTTGT	TTAGGAGAGACATTTCTAGCAA	2344	
Db	4550	ATGCTTTTCTAAATTATGCTTGAAGGCCCACTACCTTGT	TTAGGAGAGACATTTCTAGCAA	4609	
Qy	2345	AAGCAGGGGCCATTTATACACCTGAAATAGGAGAAAGGAA	CAACCCCGTTGTGTGTCNCCCTG	2404	

|||||  
4610 AAGCAGGGGCCATTATATACACCTGAACATAGAGAGAGAACACCCGGTTGTTGT -CCCTCG 4668  
QY  
2405 CTTGAGGAAGGAATTAATCTCTGAAGTCTGGGCAACAGAAAGACAAATATGGAGAGCCAAA 2464  
Db  
4669 CTTGAGGAAGGAATTAATCTCTGAAGTCTGGGCAACAGAAAGACAAATATGGACGAG -CAAA 4727  
QY  
2465 GAACTGCCGCTCTGTTCAAGTTTAAACTAAAGGATTCACCTTCCTTCCCTACCAAGGCA 2524  
Db  
4728 GAATGCCGCTCTGTTCAAGTTTAAACTAAAGGATTCACCTTCCTTCCCTACCAAGGCA 4787  
QY  
2525 GTACCCCTCTCAGACCCCAAGGGCCCAACAGGATTCCAAAAGATTTGTTAAGGACTTAAAGAC 2584  
Db  
4788 GTACCCCTCTCAGACCCCAAGGGCCCAACAGGACTTCAAAAGATTTGTTAAGGACCTTAAAGC 4847  
QY  
2585 CCAAGGCTTAGTAAACCATGATTAACCTCCCTGCAAGTAAATTCGCTAGTGGATTTGAGGAGG 2644  
Db  
4848 CCAAGGCTTAGTAAACCATGATTAACCTCCCTGCAAGTAACTCC-----AATTTTAGGAG 4900  
QY  
2645 CACAGAAACCCAGTGGACAGTGGAGGTTAGTGCAGATCTCAGATTTATCAATGGAGGC 2704  
Db  
4901 TACAGAAACCCCAAGACAGTGGAG -GGTTAGTGCAGATCTCAGATTTATCAAT -GAGGC 4958  
QY  
2705 CTTTCTCTCTTTATACCCAGCTGTACCTAGCCCTTATATCTGTGMYTTCCCAAATACCAAGA 2764  
Db  
4959 TGTGTTCTCTATATAGCCAGCTGTACCTAGCCCTTATATCTGTCTTCCCAAATACCAAGA 5018  
QY  
2765 GGAAGCAGAGTGGTTTAAACSTCTCGACCTTMAAGATGCTTCTTCTGCATCCCTGTATCA 2824  
Db  
5019 GGAAGCAGAGTGGTTTAAACSTCTCGACCTTMAAGATGCTTCTTCTGCATCCCTGTATCA 5078  
QY  
2825 TCCTGACTCTCAATTTCTGTTTGGCTTTGAGATFACCTTCAACCCACATCTCAACTCAC 2884  
Db  
5079 TCCTGACTCTCAATTTCTGTTTGGCTTTGAGATFACCTTCAACCCACATCTCAACTCAC 5138  
QY  
2885 CTGGACTRTTTTACCCCAAGGGTTTCAGGATAGYCCCATCTATTTTGGCCAGGATTAGC 2944  
Db  
5139 CTGGACTRTTTTACCCCAAGGGTTTCAGGATAGTCCCATCTATTTTGGCCAGGATTAGC 5198  
QY  
2945 CCAAGACTGAGCYARTYMCATACCTCGACACTCTTGTCTCTCTGCTAGTGGATGATTT 3004  
Db  
5199 CCAAGACTGAGCCAAATCTCATACCTCGACA -CTTGTCTCTCGTAGGTGGATGATTT 5256  
QY  
3005 ACTTTTTRGCGCYFTTCAGAAACCTTGGCCATCAAGCCACCCCAAGCCTCTTMAATTT 3064  
Db  
5257 ACTTTTGGCGGCCAATTCAGAAACCTTGGCCATCAAGCCACCCCAAGCGCTCTTCAAATTT 5316  
QY  
3065 CCTCGCYACTGTGGCTACAMGGTTTTCAAACSAARAGTCTACCTCTGCTCACAGCAGGT 3124  
Db  
5317 CCTCGCTACTGTGGCTACATGGTTTTCNAACCAAGGCTCAACTCTGCTCACAGCAGGT 5376  
QY  
3125 TAAATCTTAGGCTTAARATTTATCAAAAGGCAACCAAGGCCCTCAGTGGAGGAAYRYATCCA 3184  
Db  
5377 ---TACTTAGGGCTTAAATTTATCAAAAGGCAACCAAGGCCCTCAGTGGAGGAACATCCA 5432  
QY  
3185 GCCTATACTGGCTTATCTCTCATCYCAAAACCCCTTAAAGCACTAAGRBRRITCTTGGCT 3244  
Db  
5433 GCCTATACTGGCTTATCTCTCATCCCAAAACCCCTTAAAGCACTAAGGGGATTCCTTTGGCT 5492  
QY  
3245 AAYAGGTTTCTCCGAATGATTTCCCGAGTWTGGCRAAATAGCCAGGYCATTAWATA 3304  
Db  
5493 AATAGGTTTCTCCGAATGATTT -CCAGGATATGGGAAATAGCCAGGTCAITTAATA 5551  
QY  
3305 CASTAATTAAGGAATCTCAGAAAGCCAATACCCATTTTARTAAGATGGAYAMCTGAAGYMR 3364  
Db  
5552 CACTAATTAAGGAATCTCAGAAAGCCAATACCCATTTTARTAAGATGGACAACTGAAGTAG 5611  
QY  
3365 AAGTGGCTTTCCAGGCCCCCTAAGAGGCTTTAAACCCCAAGYCCCAAGTTTAAAGTTGCC 3424  
Db  
5612 AAGTGGCTTTCCAGGCCCC-----TAAACCCCAAGCCCCAGGTGTTAAGTTTGCC 5657  
QY  
3425 AACRGGGCAAGACTTTTSTTATVATVATCAAGAAAAAACAAGAAAYAGCTCTRGAGTCC 3484  
Db  
|||||

Db  
5658 AACAGGGCAAGACTTTTCTCATATGTTCACAG -AAAAAACAGGAATAGCTCTAGGAGTCC 5716  
QY  
3485 TTACACAGRCCCRAGGAYGAGCTTGCACACCYRTGGCBYACCTGASTRAAGAAATATGATG 3544  
Db  
5717 TTACACAGATCCGAGGATGAGCTTGCACCTGTGGCATACCTGACTAAGGAAATTTGATG 5776  
QY  
3545 TAGTGGCAAAAGGGTTGRCYTCATTTTAYVGGGTAGTGGTGGCAGTAGCAGTYKTAGTAT 3604  
Db  
5777 TAGTGGCAAAAGGGTTGACCTCATTTTAYCOGGTATGTTGGCAGTAGCAGCTTATGATAT 5836  
QY  
3605 CTGAAGCAGTTTAAATTAATACAGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGA 3664  
Db  
5837 CTGAAGCAGTTTAAATTAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTA 5896  
QY  
3665 AYRCATACCTACCTGTCTAAGGAGACTTGTGGCTGTGAGACAACYGTITTAATAATFTC 3724  
Db  
5897 ATGGCATACCTACCTGTCTAAGGAGACTTGTGGCTGTGAGACAACCTGTTTACTTAATGTC 5956  
QY  
3725 AGGCTCTATTACTTTGAARGGCCAGTGTGCRACCTGTGCACCTTGTGCAACTCTTAACCCAG 3784  
Db  
5957 AGGCTCTATTACTTTGAARGGCCAGTGTGCRACCTGTGCACCTTGTGCAACTCTTAACCCAG 6016  
QY  
3785 YCNCAITTTCTCCAGACATGAAGAAAGATPARAAYATAACTGTCAACAARTAAATTTCTC 3844  
Db  
6017 CCACATTTCTTCCAGACATGAAGAAAGATPARAAYATAACTGTCAACAARTAAATTTCTC 6076  
QY  
3845 AAACCTATGCCACTCGAGGGGACCTTGTAGAGTTCCYTTGCACTGATCCYACCTTCAAC 3904  
Db  
6077 AAACCTATGCCACTCGAGGGGACCTTTTAGAGGTTCCYTTGCACTGATCCCGACC -TCAAC 6135  
QY  
3905 TTGTATATCTGATGGAAGTTCTTTGTAGAAAAAGGACTTCGAAAAGYGGGTATGACGTG 3964  
Db  
6136 TTGTATATCTGATGGAAGTTCTTTGTAGAAAAAGGACTTCGAAAAGTGGGGTATGACGTG 6195  
QY  
3965 GTCAGTGATATGGAATAYTTGAAAGTAATCCCTCTCACTCCAGGAACCTAGTGTGCTVAGCTR 4024  
Db  
6196 GTCAGTGATATGGAATAYTTGAAAGTAATCCCTCTCACTCCAGGAACCTAGTGTGCTVAGCTR 6255  
QY  
4025 GCAGAACTAATATAGCCYTCATYKGGGCATAGAAATTTAGGAGAAAGRAAAAGGGYAAATATA 4084  
Db  
6256 GCAGAACTAATATAGCCCTCACTTGGGCATAGAAATTTAGGAGAAAGRAAAAGGGYAAATATA 6315  
QY  
4085 TATACAGACTCTARTATGCTYACCTAGTCTCATGCTCCCATGMRGCAATATGSAAGAA 4144  
Db  
6316 TATACAGACTCTAATATGCTTACTAGTCTCTCATGCTCCCATGCAAGCAATATGGAAGAA 6375  
QY  
4145 AGGGAATTTCTTAACCTTCYAGRGAAACCTATCAMACATCAGGAAGCCATTTAGGARATTA 4204  
Db  
6376 AGGGAATTTCTTAACCTTCYAGRGAAACCTATCAMACATCAGGAAGCCATTTAGGARATTA 6435  
QY  
4205 TTATYGGGWTGACAGAAACCTBARAGAGTGGMAGTCTTACACTGCTGCTGGGTCTATCANAA 4264  
Db  
6436 TTATYGGGWTGACAGAAACCTBARAGAGTGGMAGTCTTACACTGCTGCTGGGTCTATCANAA 6495  
QY  
4265 GGAAGRAAAAGGGAAATASAAAGAAATGCAAGCAKATATTTGAAGCAAAAGAGCTGCA 4324  
Db  
6496 GGAAGRAAAAGGGAAATASAAAGAAATGCAAGCAKATATTTGAAGCAAAAGAGCTGCA 6555  
QY  
4325 AGGCAAGACCTCCATAGGAATGCTTATTAACCTTCCCTTTAGTATAGGGTAATCCCTTC 4384  
Db  
6556 AGGCAAGACCTCCATAGGAATGCTTATTAACCTTCCCTTTAGTATAGGGTAATCCCTTC 6615  
QY  
4385 CGGGAACCAAGCCCGCTAGTACTCAGCAGGAGAAACAGAAATGGGAACCTCTCAGGAG -CAG 4443  
Db  
6616 CGGGAACCAAGCCCGCTAGTACTCAGCAGGAGAAACAGAAATGGGAACCTCTCAGGAGAG 6675  
QY  
4444 TTTTCTCCCTCGGGACGTTAGCAGCTTGAAGAGGGGAAATACCTTTTTCCTGCAACTAT 4503  
Db  
6676 TTTTCTCCCTCGGGACGTTAGCAGCTTGAAGAGGGGAAATACCTTTTTCCTGCAACTAT 6735  
QY  
4504 CCAATGGAAATTAATAAAACCTTCAATCAAAACCTTTCACTTAGGCATCGATAGCACCA 4563  
Db  
6736 CCAATGGAAATTAATAAAACCTTCAATCAAAACCTTTCACTTAGGCATCGATAGCACCA 6795

QY	4564	TCARATGCGCCAAATCACTATTATTTACTGGACGAGGCGCTTTTCAAAACTATCAAGCARATAKT	4623	5644	CCCCCTCCATGCCGCTGTATGACCAAGTAGCTCCCTCTYVACMAGAGTTTCTATGGAGAAATG	5700
DB	6796	TCAGATGGCCAAATCAITTTATTTACTGGACCGAGGCGCTTTTCAAAACTATCAAGCAGATAGT	6855	7875	CCCCCTCCATGCCGCTGTATGACCAAGTAGCTCCCTCTYVACMAGAGTTTCTATGGAGAAATG	7934
QY	4624	CAGGSCCTGTGAATGTGCCARABAAATAAATCCCTGCTVATCGCCAAAGCTCTCTCAGG	4683	5704	CAGCGTCCCGGAAATATTCATGATGCCCATCGTATAGGAGTCTTTTCTAAGGAAACCCCCACC	5763
DB	6856	CAGGSCCTGTGAAGTGTGCCAGAGAAATAATCCCTGCTVATCGCCAAAGCTCTCTCAGG	6915	7935	CAGCGTCCCGGAAATATTTGATGCCCCATCGTATAGGAGTCTTTTCTAAGGAAACCCCCACC	7994
QY	4684	APAAACAARAACAGGCCATTTACCTGTRABARAATCTGGCAACTGATTTTACCACAAGCCC	4743	5764	TTCACTGCCCAACCCATATGCCCCCAACTGCTATCACTCTGCCACTCTTTTGCATGCAAT	5823
DB	6916	AGAAACAAGAACAGGCCATTTACCTGGAAGAGACTGGCAACTGATTTTACCACAAGCCC	6975	7995	TTCACTGCCCAACCCATATGCCCCCAACTGCTATCACTCTGCCACTCTTTTGCATGCAAT	8054
QY	4744	AAACCTCAGGATTTTCAGTATCTACTAGTCTGGGTARATCTTTACGCGGTGGGCGARAG	4803	5824	GCAAACTACTCATTTTGGACAGGAAAAATGATTAATCTCTAGTTGTCTCGAGGACTTGA	5883
DB	6976	AAACCTCAGGATTTTCAGTATCTACTAGTCTGGGTARATCTTTACGCGGTGGGCGARAG	7035	8055	GCAAACTACTCATTTTGGACAGGAAAAATGATTAATCTCTAGTTGTCTCGAGGACTTGA	8114
QY	4804	GCCTTCCCTCTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCATGAATA	4863	5884	GTCACTGTCTGTGGACTTACTTACCACAACTGGGTATCTCTGATGGGGGTGGAGTTCAA	5943
DB	7036	GCCTTCCCTCTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCATGAATA	7095	8115	GTCACTGTCTGTGGACTTACTTACCACAACTGGGTATCTCTGATGGGGGTGGAGTTCAA	8174
QY	4864	ATTTCCAGATTTCCGACTTCCCGAGGCTTACAGAGTGACATAGCCCTGTCTCCAGGCC	4923	5944	GATCAGGCAAGGAAAAACATGTAAAGAAAGTAAATCTCCAACTACATGAACCCCTCGTACCCAT	6003
DB	7096	ATTTCCAGATTTCCGACTTCCCGAGGCTTACAGAGTGACATAGCCCTGTCTCCAGGCC	7155	8175	GATCAGGCAAGGAAAAACATGTAAAGAAAGTAAATCTCCAACTACATGAACCCCTCGTACCCAT	8234
QY	4924	ACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGA	4983	6004	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCGTACCCAT	6063
DB	7156	ACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGA	7215	8235	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCGTACCCAT	8294
QY	4984	AGGCCACAGTCTCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAGAGACATCTAAAA	5043	6064	ACTCGCTCGTAAAGCCTATTTAATACCACTTACCACTTGGGCTCATGAGGTCTCGGCCCCAA	6123
DB	7216	AGGCCACAGTCTCTCAGGGAAGGTCGAGAAAATGAATGAAAYACTCAAGAGACATCTAAAA	7275	8295	ACTCGCTCGTAAAGCCTATTTAATACCACTTACCACTTGGGCTCATGAGGTCTCGGCCCCAA	8354
QY	5044	AAGCAAAACCCAGGAAAACCCACTCAATGGCTGTCTGTGGCTATAGCTTTAAAAAGA	5103	6124	AACCTCTACTAACTGTGTGGATATGCCTCCCTGAACTTCARGCCATATGTGTTCAATCCCT	6183
DB	7276	AAGCAAAACCCAGGAAAACCCACTCAATGGCTGTCTGTGGCTATAGCTTTAAAAAGA	7335	8355	AACCTCTACTAACTGTGTGGATATGCCTCCCTGAACTTCARGCCATATGTGTTCAATCCCT	8414
QY	5104	ATCTGCAACTTTCCCAAAAGCAGGACTTAGCCCATACGAATGCTGTATGGAGGCC	5163	6184	GTACTGTAAACATGGAAACAACTTCAGCACAGAAAATAAACACACTTCCTGTTTGTAGTGA	6243
DB	7336	ATCTGCAACTTTCCCAAAAGCAGGACTTAGCCCATACGAATGCTGTATGGAGGCC	7395	8415	GTACTGTAAACATGGAAACAACTTCAGCACAGAAAATAAACACACTTCCTGTTTGTAGTGA	8474
QY	5164	TTCAATAACCAATGACTTGTCTGTGACCCAGACAGCAACTTGTGTGACAGATCACCT	5223	6244	CCTCTGTGTTTCCAAATGGAATTAACCATCTCAACCTCACTGCTGTGTAATTTTATAG	6303
DB	7396	TTCAATAACCAATGACTTGTCTGTGACCCAGACAGCAACTTGTGTGACAGATCACCT	7455	8475	CCTCTGTGTTTCCAAATGGAATTAACCATCTCAACCTCACTGCTGTGTAATTTTATAG	8534
QY	5224	CCTTAGCCAAATATCAACAAGTTCTTAAAAATTTACAGGAACCTATCCCTGAGAGAGG	5283	6304	AATATCATACACAAACAACTCCCAATGCATCAGTGGGTAACTCCTCCACACACAAATA	6363
DB	7456	CCTTAGCCAAATATCAACAAGTTCTTAAAAATTTACAGGAACCTATCCCTGAGAGAGG	7515	8535	AATATCATACACAAACAACTCCCAATGCATCAGTGGGTAACTCCTCCACACACAAATA	8594
QY	5284	GAAAGAACTTATCCACCCWGTGACATGTTATGTTCAAGTCCCTTCCTAAATTCCT	5343	6364	GTCTGCTTACCTCAGGAATATTTTTTGTCTGTGTGTACCTCAGCCCTATCGTTTGTGAAT	6423
DB	7516	GAAAGAACTTATTCACCTTGTGACATGTTATGTTCAAGTCCCTTCCTAAATTCCT	7575	8595	GTCTGCTTACCTCAGGAATATTTTTTGTCTGTGTGTACCTCAGCCCTATCGTTTGTGAAT	8654
QY	5344	CATCCCTAGATATCATCTCTGGAGAGCCCTACCCAGTCACTTTATYTACCCCACTGCGG	5403	6424	GGCTCTTCAGATCTATGTGCTTCTCTCATCTTTAGTGCCCTCATGBCCATCTACACT	6483
DB	7576	CATCCCTAGATATCATCTCTGGAGAGCCCTACCCAGTCACTTTATYTACCCCACTGCGG	7635	8655	GGCTCTTCAGATCTATGTGCTTCTCTCATCTTTAGTGCCCTCATGBCCATCTACACT	8714
QY	5404	TTAAAGTGGCTGGAGTGAGTCTTTGGATATACATCACTTGAGTCAAACTCTGGATATGCG	5463	6484	GAAACAGATTTTATACAGTTATGTCTATCTTAAGCCCGCAACAAAGAGTACCCATCTCT	6543
DB	7636	TTAAAGTGGCTGGAGTGAGTCTTTGGATATACATCACTTGAGTCAAACTCTGGATATGCG	7695	8715	GAAACAGATTTTATACAGTTATGTCTATCTTAAGCCCGCAACAAAGAGTACCCATCTCT	8774
QY	5464	CAAAAGGAACTTGAAATTCAGGAGAACCGCTAGCTATTCCTGTGAACCTCTTAGAGGAT	5523	6544	CCTTTTGTATAGGACGAGAGTGTAGGTGCACTAGGTACTGCGCATTTGGCGGTATACA	6603
DB	7696	CAAAAGGAACTTGAAATTCAGGAGAACCGCTAGCTATTCCTGTGAACCTCTTAGAGGAT	7755	8775	CCTTTTGTATAGGACGAGAGTGTAGGTGCACTAGGTACTGCGCATTTGGCGGTATACA	8834
QY	5524	TGCGCCTGCTCTTCAAAACAACACGAGGAAAGTAACTAAATCATTAATCCCCCATG	5583	6604	ACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTTAATGGGGACATGGAACGGGTC	6663
DB	7756	TGCGCCTGCTCTTCAAAACAACACGAGGAAAGTAACTAAATCATTAATCCCATG	7814	8835	ACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTTAATGGGGACATGGAACGGGTC	8894
QY	5584	GSCCTCCCTTATCATATTTTCTCTKATGTTSTTTTACCCCTSTTTTCACTCTCACTGCA	5643	6664	GCGGACTCCTCGGTCAACTTGGCAAGTCAACTTAACTCCCTAGCAGCAGTAGTCTCTCRA	6723



```
|||||
8955 AATCGAAGAGCTTTAGACTGCTAAACGCTGAAGAGGGGAACCTGTTTATTTTAGGG 9014
QY GAAGAATGCTGTTATTTATGTTTAAATCAATCCGAATCGTCACTGAGAAAGTTTAAAGAAAT 6843
Db GAAGAATGCTGTTATTTATGTTTAAATCAATCCGAATCGTCACTGAGAAAGTTTAAAGAAAT 9074
QY CSAGATCGAATAACAAGTAKAGCAGAGAGCTTCCGAACACTGGACCTGGGGCCTCCTC 6903
Db CGAGATCGAATAACAAGTAKAGCAGAGAGCTTCCGAACACTGGACCTGGGGCCTCCTC 9134
QY AGCCATGATGATCCCTCGATTCCTCCCTCTTAGGACCTCTAGCAGCTATATATTTGCTA 6963
Db AGCCATGATGATCCCTCGATTCCTCCCTCTTAGGACCTCTAGCAGCTATATATTTGCTA 9194
QY CTCCTCTTTGGACCTGATCTATCTTTTACCTCTCTTTTAACTTTTGTCTCTTCAGAAATCGAA 7023
Db CTCCTCTTTGGACCTGATCTATCTTTTAACTCTCTTTTAACTTTTGTCTCTTCAGAAATCGAA 9254
QY GCTGTAAACTACAATGGAGCCCAAGATGCACTCCAGACTTAAGACTTACCGCAGACCC 7083
Db GCTGTAAACTACAATGGAGCCCAAGATGCACTCCAGACTTAAAGACTTACCGCAGACCC 9314
QY CTGACCGGCTGTYTAGCCCAAGATCTGATGTTTAAATGACATCAAAAGGCAACCCCTCTCGAG 7143
Db CTGACCGGCTGCTAGCCCAAGATCTGATGTTTAAATGACATCAAAAGGCAACCCCTCTCGAG 9374
QY GAAATCTCAGTGCACAACCTCTACTACGCCCAATTCAGCAGGAGAGAGTTAGAGCGGT 7203
Db GAAATCTCAGTGCACAACCTCTACTACGCCCAATTCAGCAGGAGAGAGTTAGAGCGGT 9434
QY SCTCGGCCCAACCTCCCAACAGCAGCTTAGGTTTCTGTTGAGATGGGGGACTGAGAGAC 7263
Db SCTCGGCCCAACCTCCCAACAGCAGCTTAGGTTTCTGTTGAGATGGGGGACTGAGAGAC 9494
QY AGCACTAGCTGATTTCTTAGGCTGATTAAGAACTCYTAAGCTTAGSTGGGAAGGTGACC 7323
Db AGCACTAGCTGATTTCTTAGGCTGATTAAGAACTCYTAAGCTTAGSTGGGAAGGTGACC 9554
QY ACATCCACCTTTAAACAGCGGCTTTGCAACTTAGTGTACACCTGACCAATCAAGAGAGAGCTC 7383
Db ACATCCACCTTTAAACAGCGGCTTTGCAACTTAGTGTACACCTGACCAATCAAGAGAGCTC 9614
QY ACTAAATGCTAATTTAGGCAAGAGAGAGGCTTAAAGAAATAGCCCAATCATTTATTCGWTG 7443
Db ACTAAATGCTAATTTAGGCAAGAGAGAGGCTTAAAGAAATAGCCCAATCATTTATTCGWTG 9674
QY AGAGCAGCAGCAGGAGGCAATGATCGGATATAAACCCAAAGTCTTCGAGCCGCAACGG 7503
Db AGAGCAGCAGCAGGAGGCAATGATCGGATATAAACCCAAAGTCTTCGAGCCGCAACGG 9734
QY CAACCCCTTTGGGTCCTCCCTCTTGTATGGAGCTCTGTTTTCATGCTATTTCACTCT 7563
Db CAACCCCTTTGGGTCCTCCCTCTTGTATGGAGCTCTGTTTTCATGCTATTTCACTCT 9794
QY ATTAAATCTTGCACTGCR 7582
Db ATTAAATCTTGCACTGCA 9813
```

```
RESULT 11
AY101584
LOCUS Homo sapiens isolate 10222 bp DNA linear. PRI 11-FEB-2004
DEFINITION Homo sapiens isolate 79 endogenous retrovirus HERV-W, ERVWE1 locus,
allele A, complete sequence.
ACCESSION AY101584
VERSION AY101584.1 GI:37544403
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
```

```
REFERENCE
AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
TITLE The endogenous retroviral locus ERVWE1 is a bona fide gene involved
in hominoid placental physiology
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
PUBMED 14757826
REFERENCE
AUTHORS Mallet,F., Bouton,O. and Oriol,G.
TITLE Direct Submission
JOURNAL Submitted (06-MAY-2002) Retrovirology Department, UMR 2142
CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
FEATURES
source
1..10222
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="79"
/isolation_source="PBMC of individual 79"
/db_xref="Taxon:9606"
/chromosome="7"
/map="7q21-q22"
/sex="male"
/cell_type="PBMC"
/note="Caucasian"
isolated by PCR; ERVWE1 locus, allele A
endogenous_virus: HERV-W
1..780
join(248..811,2887..9813)
/note="putative mRNA transcript 1; gag-pol pseudogene"
join(248..811,7585..9813)
/product="envelope glycoprotein"
/note="putative mRNA transcript 2"
join(248..811,9248..9813)
/note="putative mRNA transcript 3; no coding region"
7812..9428
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAM68163.1"
/db_xref="GI:37544404"
/translaton="MALPYHIFLFTVLLPSFTLAPPCCMTSSSPYQEFILWRMORP
TNDVSTQTMGSDGGVQDOAREKHVKVLSOLTQVHGTSSPYKGLDLSKLHETLRT
GIVDAPYRSLSGKPTPTAHTHMPNCHSATHLCMHANTHYWTKMINPSPCGGLGV
TNCVYTGTCGSDGGVQDOAREKHVKVLSOLTQVHGTSSPYKGLDLSKLHETLRT
HTRVLSFNTLTGLHEVSAQNPTNCWICLPLNRPYVSIIPVEQNNPSEINTSV
LVGLPVSNEIHTSNTLVKFSNTTYTNSQCRWVPTQIVCLPSGIFPVCSTSA
YRCLNGSESCFLSPVPMIYTEDLYNYSKPRKRVPLPFVIGAGVGLGALG
TGIGITSTQFYKLSQELNGDMERVADSLVTLQDQNSLAAVVLQNRALDLTAE
RGGTCLFLGBCCVVNSQGIIVTEKVEIRDIORRABELRNTGWLISOWMPWILP
FLGPLAAILILLFLPGICFNLLVNFVSRIRIEAVKLQMEPKMQSKTKIYRRPLDRPAS
RSDVNDKNGTPPEISAAQPLLRPNSSGS"
9487..10222
LTR
ORIGIN
Query Match 86.0%; Score 6440; DB 5; Length 10222;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 6676; Conservative 195; Mismatches 55; Indels 53; Gaps 27;
QY 606 GATGGGAAACGTTCCCGCAAGACAAAACGCCCTTAAGACGTATTCGGAATAATGGGA 665
Db 2886 GATGGGAAACGTTCCCGCAAGACAAAACGCCCTTAAGACGTATTCGGAATAATGGGA 2945
QY 666 MCAATTTGACCTCAGACACTAAGAAAGAGCTTATATCTTCGAGTGGCCGCTG 725
Db 2946 CCAATTTGACCTCAGACACTAAGAAAGAGCTTATATCTTCGAGTGGCCGCTG 3005
QY 726 GCATCTCTGAGGAGTATAAATTATACACCATCTTACAGTACACTCTTTCTGTAGAA 785
Db 3006 GCATCTCTGAGGAGTATAAATTATACACCATCTTACAGTACACTCTTTCTGTAGAA 3065
QY 786 AAGGCAATGAGTGAACTGCATAGTACAAACTTTCTTTTCAATTAAGAGACACTCAC 845
Db 3066 AAGGCAATGAGTGAACTGCATAGTACAAACTTTCTTTTCAATTAAGAGACACTCAC 3125
```





Db 5257 ACTTTTGGCGCCCATTCAGAAACCTTGTGCCATCAGCCACCCCAAGTGTCTCTCAATTT 5316  
QY 3065 CCTCGCYACCTGTGTGCTCAGNGTTTCAAACSAAPARGCTCARTCTCTCTCAACAGAGGT 3124  
Db 5317 CCTCGCTACCTGTGTGCTACATGTTTCAAACCAAAGGCTCAACTCTCTCTCACAGCAGGT 5376  
QY 3125 TAATACTTAGGRCCTAARATTTCAAAGGACCAAGGCCCTCAGTGAAGGAAYATCCA 3184  
Db 5377 ----TACTTAGGGCTAATAATTTATCAAAGGCCAACAGGGCCCTCAGTGAAGGAACATCCA 5432  
QY 3185 GCCTATACCTGGCTTATCTCTCATCYCAAAACCCCTAAAGCACTAAAGRRRTTCTTTGGGCT 3244  
Db 5433 GCCTATACCTGGCTTATCTCTCATCCCAAACCCCTAAAGCACTAAAGGGGRTTCTTTGGGCT 5492  
QY 3245 AAYAGGYTCTGCCGAATATGGAATCTCCAGGTWTGGCRAAATAGCCAGGAYCATPAMATA 3304  
Db 5493 AATAGGTTTTCGCCGAATAATGGAAT - CCCAGGTATGGCGAATAGATGCCAGGTCAITTAATA 5551  
QY 3305 CASTAATTAAGGAACCTCAGAAAGCAATACCCATTTARTAGATGGAYMCTGAGYMR 3364  
Db 5552 CACTAATTAAGGAACCTCAGAAAGCAATACCCATTTAGTAGATGGCAACTGAAGTAG 5611  
QY 3365 AAGTGGCTTTCCAGGCCCTTAAAGAGGCTTTAAACCCCAAGYCCAGTGTTAAGYTTGCC 3424  
Db 5612 AAGTGGCTTTCCAGGCC-----TAAACCCAGGCCAGTGTTAAGYTTGCC 5657  
QY 3425 AACRGCGCAAGACTTTTSTTYATYRTCAAGAAAAAAACAGRAAYAGCTCTRGAGTCC 3484  
Db 5658 AACAGGCAAGACTTTTCTTCATATGTACAG - AAAAAACAGGAATAGCTCTAGGAGTCC 5716  
QY 3485 TTACACAGTCCRAGGAYGAGCTTGCAACCVRTGGCRYACCTGASTAAGGAAATGTATG 3544  
Db 5717 TTACACAGATCCGAGGGATGAGCTTGCACCTGTGGCACTACCTGACTAAGGAAATGTATG 5776  
QY 3545 TAGTGGCAAGGGTGRCTCATTTGTTAYGGGTAGTGTGGCAGTACGAGYKTAGTAT 3604  
Db 5777 TAGTGGCAAGGGTGTACCTCAITGTTTACGGGTAGTGTGGCAGTACGAGTCTTAGTAT 5836  
QY 3605 CTGAAGCAGTTAAAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGA 3664  
Db 5837 CTGAAGCAGTTAAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGA 5896  
QY 3665 AYRGCACTACTCTGCTAAGGAGACTTGTGTGCTGTGAGCAACYGTGTTACTTAATATTC 3724  
Db 5897 ATGGCATACTCACTGCTAAGGAGACTTGTGTGCTGTGAGCAACTGTGTTACTTAATGTCTC 5956  
QY 3725 AGGCTCTATTACTGAAGGCCAGTGTGCRACCTGTGCACTGTGCAACTCTTAAACCCAG 3784  
Db 5957 AGGCTCTATTACTGAAGGCCAGTGTGCRACCTGTGCAACTGTGCAACTCTTAAACCCAG 6016  
QY 3785 YGNCATTTTCTCCAGACATGAAGAAAAAGATARAAYATACTGTCAACAARTAAATTTCTC 3844  
Db 6017 CCACATTTCTCCAGACATGAAGAAAAAGATAAACAATACTGTCAACAAGTAATTTCTC 6076  
QY 3845 AAAACCTATGCCACTCGAGGGGACCTTYTAGARGTTCVYTGTACTGATTCYGAACCTTCAAC 3904  
Db 6077 AAAACCTATGCCACTCGAGGGGACCTTTTAGAGGTTCTTTGACTGATCCCGACC - TCAAC 6135  
QY 3905 TTGTATACTGATGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAAGYGGGGTATCGATG 3964  
Db 6136 TTGTATACTGATGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAAGTGGGGTATCGATG 6195  
QY 3965 GTCAGTGATAATGGAATAYTTGAAAGTAAATCCCTCACTCCAGGAACCTAGTCTYAGCTR 4024  
Db 6196 GTCAGTGATAATGGAATACTTGAAGTAATCCCTCACTCCAGGAACCTAGTCTCAGCTA 6255  
QY 4025 GCAGAACTAATAGCCYTCATYTKGGGCACTAGAAATTAGGAGAGRAAAAAAGGGYAAATATA 4084  
Db 6256 GCAGAACTAATAGCCCTCACTTGGGCACCTAGAAATTAGGAGAGRAAAAAAGGGCAATATA 6315  
QY 4085 TATACAGACTCTPRARTACTCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSRAGAA 4144

Db 6316 TATACAGACTCTAAATATGCTTACTAGTCTCTCATGCCCATCGACCAATATGGAAGAA 6375  
QY 4145 AGGGAATTCCTAACTTCYAGRGAAACACCTATCAMAATCAGGAAGCCATTTAGGARATTA 4204  
Db 6376 AGGGAATTCCTAACTTCYAGRGAAACACCTATCAMAATCAGGAAGCCATTTAGGAATTA 6435  
QY 4205 TTAATGCGWGTACAGAAACCTTARAGAGTGTGMACTCTTACACTGCTCYGGGTTCATCANAAA 4264  
Db 6436 TTAATGCGWGTACAGAAACCTTAAAGAGTGTGMACTCTTACACTGCGGGGTTCATCAGAAA 6495  
QY 4265 GGAAGRAAAGGGGAAATATAAAGRAATGTCACACAKATATTTGAAGCMAAAGAGCTGCA 4324  
Db 6496 GGAAGRAAAGGGGAAATATAAAGRAATGTCACACAGATATTTGAAGCCAAAGAGCTGCA 6555  
QY 4325 AGGCAGAGCCCTCATTTAGAAATGCTTTTAAACTTCCCTTAGTATATAGGGTAATCCCTTC 4384  
Db 6556 AGGCAGAGCCCTCATTTAGAAATGCTTTTAAACAACCCCTTAGTATAGGGTAATCCCTTC 6615  
QY 4385 CGGGAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAAATGGGAACTCTCACGAGG - CAG 4443  
Db 6616 CGGGAACCAAGCCCCAGTACTCAGCAGGAGAAACAGAAATGGGAACTCTCACGAGGACAG 6675  
QY 4444 TTTTCTCCCTCGGGACGGTTAGCCACTGGAAGAGGGGAAATATCTTTTGCCTGCAACTAT 4503  
Db 6676 TTTTCTCCCTCGGGACGGCTAGCCACTGGAAGAGGGGAAATATCTTTTGCCTGCAACTAT 6735  
QY 4504 CCAATGGAATATCTTAAACCTTTCATAAACCTTTTCACTTAGGCATCGATAGCACCA 4563  
Db 6736 CCAATGGAATATCTTAAACCTTTCATAAACCTTTTCACTTAGGCATCGATAGCACCA 6795  
QY 4564 TCARATGCCAAATCATTTATTTAGTGGACCAAGGCTTTTCAAAAATCAAGCAGATAT 4623  
Db 6796 TCARATGCCAAATCATTTATTTAGTGGACCAAGGCTTTTCAAAAATCAAGCAGATAT 6855  
QY 4624 CAGGCTGTGAATGTGCCARAAAAATAATCCCTGCTTATGCCAAGCTCTTTCAGG 4683  
Db 6856 CAGGCTGTGAATGTGCCAGAGAAATAATCCCTGCTTATGCCAAGCTCTTTCAGG 6915  
QY 4684 ARAACAAARAACAGGCCATTAACCTGARAARAACCTGGAACATGATTTTACCAAGGCC 4743  
Db 6916 AGAACAAARAACAGGCCATTAACCTGGAAGAAGACTGGCAACTGATTTTACCAAGGCC 6975  
QY 4744 AAACTCAGGATTTCACTATCTAGTCTGGGTABATATCTTTCACCGGTGGGCARAG 4803  
Db 6976 AAACTCAGGATTTCACTATCTAGTCTGGGTABATATCTTTCACCGGTGGGCARAG 7035  
QY 4804 GCCTTCCCTCTAGGACAGAAAAAGGCCCAAGAGTAATAAAGGCACTAGTTTCATGAATA 4863  
Db 7036 GCCTTCCCTCTAGGACAGAAAAAGGCCCAAGAGTAATAAAGGCACTAGTTTCATGAATA 7095  
QY 4864 ATTCAGATTCGAGCTTCCCGAGGCTTACAGAGTGA CAATAGCCCTGCTTTCAGGCC 4923  
Db 7096 ATTCAGATTCGAGCTTCCCGAGGCTTACAGAGTGA CAATAGCCCTGCTTTCAGGCC 7155  
QY 4924 ACAGTACCCAGGAGTATCCAGCGTTAGGTATACCATATCACTTACACTGCGCCCTGA 4983  
Db 7156 ACAGTACCCAGGAGTATCCAGCGTTAGGTATACCATATCACTTACACTGCGCCCTGA 7215  
QY 4984 AGGCCACAGTCTCAGGGAGGCTCGAGAAAAATGAATGAAAYACTCAAAAGGACATCTAAAA 5043  
Db 7216 AGGCCACAGTCTCAGGGAGGCTCGAGAAAAATGAATGAAACACTCAAAAGGACATCTAAAA 7275  
QY 5044 AAGCAAAACCCAGGAAACCCACCTCACATGCGCTGTCTGTGCTTATAGCTTAAAAAGA 5103  
Db 7276 AAGCAAAACCCAGGAAACCCACCTCACATGCGCTGTCTGTGCTTATAGCTTAAAAAGA 7335  
QY 5104 ATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATCTGTATGAGAGGCC 5163  
Db 7336 ATCTGCAACTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATCTGTATGAGAGGCC 7395  
QY 5164 TTCAATAACCAATGACCTTGTGCTTGACCCAAAGACAGCAACTTAGTTGACAGATCACCT 5223  
Db 7396 TTCAATAACCAATGACCTTGTGCTTGACCCAAAGACAGCAACTTAGTTGACAGATCACCT 7455



```
|||||
Db 9615 ACTAAATGCTAATTAGCAAAACACAGAGGTGAAGAAATAGCCAAATCATCTATTGCGCTG 9674
|||||
Qy 7444 AGAGCACAGCAGGAGGACAATATGATCGGATATATAACCCAAAGTGTTCAGCGCGGCAACGG 7503
|||||
Db 9675 AGAGCACAGCAGGAGGACAATATGATCGGATATATAACCCAAAGTGTTCAGCGCGGCAACGG 9734
|||||
Qy 7504 CAACCCCTTTGGGTCGCCCTCCCTTTGATCGGAGCTCTGTTCATGCTATTTCACCTCT 7563
|||||
Db 9735 CAACCCCTTTGGGTCGCCCTCCCTTTGATCGGAGCTCTGTTCATGCTATTTCACCTCT 9794
|||||
Qy 7564 ATTAATCTTGCARCTGCR 7582
|||||
Db 9795 ATTAATCTTGCARCTGCA 9813
|||||

RESULT 12
BD221808 10499 bp DNA linear PAT 17-JUL-2003
LOCUS BD221808
DEFINITION Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses.
ACCESSION BD221808
VERSION BD221808.1 GI:33031578
KEYWORDS JP 2002518051-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homniidae; Homo.
REFERENCE 1 (bases 1 to 10499)
AUTHORS Alliel,P.M., Perin,J.P. and Rieger,F.
TITLE Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses
JOURNAL Patent: JP 2002518051-A 3 25-JUN-2002;
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE
COMMENT OS Homo sapiens (human)
PN JP 2002518051-A/3
PD 25-JUN-2002
PF 23-JUN-1999 JP 2000556036
PI PATRICK M ALLIEL, JEAN PIERRE PERIN, FRANCOIS RIEGER PC
C12N15/09, A01K67/027, A61K31/711, A61K39/21, A61K48/00, A61P21/00, PC
A61P25/00
PC A61P37/06, C07K14/15, C12Q1/68, C12Q1/70, C12N15/00 CC Nucleic
sequence and deduced protein sequence family with CC
human endogenous
CC retroviral motifs, and their uses
FH Key Location/Qualifiers
FT source 1..10499
FT /organism='Homo sapiens (human)'.
FEATURES
source
1..10499
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

ORIGIN
Query Match 85.9%; Score 6436; DB 2; Length 10499;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 6680; Conservative 195; Mismatches 50; Indels 54; Gaps 28;

Qy 606 GATGGAAACGTTCCCGCAAGACAAAACCGCCCTAAGACGTATTCTGGAPAAATTGGGA 665
Db 2953 GATGGAAACGTTCCCGCAAGACAAAACCGCCCTAAGACGTATTCTGGAGAAATTGGGA 3012
Qy 666 MCAATTTGACCTCAGACACTAAGAAAGAAACGACTTATTCTTCAGTGCAGTCCGCTG 725
Db 3013 CCAATTTGACCTCAGACACTAAGAAAGAAACGACTTATTCTTCAGTGCAGTCCGCTG 3072
Qy 726 GCACCTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACACTCTTTTGTAGAA 785
Db 3073 GCACCTCCTGAGGGAAGTATAAATTATAACACCATCTTACAGCTAGACCTCTTTGTAGAA 3132
```

```
Qy 786 AAGCAAAATGAGTGAAGTGCATATAAGTACAAACTTTCTTTTCATTATAGACAACTCAC 845
Db 3133 AAGCAAAATGAGTGAAGTGCATATAAGTACAAACTTTCTTTTCATTATAGACAACTCAC 3192
Qy 846 AATTATGTAAAAAGTGTGATTTATGCGCTACAGAAAGCCTTCAGAGTCTACCTCCCTATC 905
Db 3193 AATTATGTAAAAAGTGTGATTTATGCGCTACAGAAAGCCTTCAGAGTCTACCTCCCTATC 3252
Qy 906 CCAGCATCCCCGACTCCTTCCCAATTAATTAAGAGACCCCTTCAACCAATGGTCCAA 965
Db 3253 CCAGCATCCCCGACTCCTTCCCAACTAATTAAGAGACCCCTTCAACCAATGGTCCAA 3312
Qy 966 AAGGAGATAGACAAAAGGTAAACAGTGAACCAAGAGTGCATATTATCCCAATATAGA 1025
Db 3313 AAGGAGATAGACAAAAGGTAAACAGTGAACCAAGAGTGCATATTATCCCAATATAGA 3372
Qy 1026 CCCCTCCCAAGCAGTGGGAGGAAGAAATTCGGCCCAAGCAGAGTGCATGTGCTTTTYY 1085
Db 3373 CCCCT-CCAAGCAGTGGGAGGAAGAAATTCGGCCCAAGCAGAGTGCATGTGCTTTTTC 3431
Qy 1086 TCTCCAGACTTAAGCAAAATAAACAAGACTTAGTAAATTTCTCAGATAAYCCCTGATGG 1145
Db 3432 TCTCCAGACTTAAGCAAAATAAACAAGACTTAGTAAATTTCTCAGATAAYCCCTGATGG 3491
Qy 1146 CTATATTGRTGTTTAAACAAGGTTAGGACAATTTCTTGATCTGCATCGAGAGATATATA 1205
Db 3492 CTATATTGATGTTTAAACAAGGTTAGGACAATTTCTTGATCTGCATCGAGAGATATA-A 3550
Qy 1206 TGTCACTGCTAAATCAGACACTAAACCCCAATGAGAGAAGTGCCCAACATAACTGCAGCT 1265
Db 3551 TGTCACTGCTAAATCAGACACTAAACCCCAATGAGAGAAGTGCCCAACATAACTGCAGCT 3610
Qy 1266 GAGGTTTGGCGATCTCTGGTATCTCAGTCAAGTCAATGATANGGATGACACAGAAGG 1325
Db 3611 GAGGTTTGGCGATCTCTGGTATCTCAGTCAAGTCAATG--ATAGGATGACACAG-AGG 3667
Qy 1326 AAAGANAATGATCCCAAGCAGCAGCAGTTCAGTTCAGTCAAGTTCAGTTCAGTTCAGT 1385
Db 3668 AAAGANAATGATCCCAAGCAGCAGCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGT 3724
Qy 1386 ACAGAAATCAGTAAACATGGAGATTTGTTGTCGACACATTTGCTAACTTGTGTGTASNA 1445
Db 3725 ACACAGAATCAGAAACAT-GGAGATTTGTTGTCGACACATTTGCTAACTTGTGTGTAGAA 3783
Qy 1446 GGACTAAGAAAACTASGAAAGAAATCTAYGAATTTACTCAATGATGTCACATACACA 1505
Db 3784 GGACTAAGAAAACTAGGAAG-AGAGTCTATGAATTTACTCAATGATGTCACATACACA 3842
Qy 1506 GGGGAAGGGAAGAAAATCCTACTGCTTCTGAGAGACTAAGGAGGAGCATGAGGAAGC 1565
Db 3843 -GGGGAAGGGAAGAAAATCCTACTGCTTCTGAGAGACTAAGGAGGAGCATGAGGAAGC 3901
Qy 1566 GTGCTCTCTGTCACCTGACTCTTCTGAAGGCAACTAATCTTAAGCGTAAAGTTTATCA 1625
Db 3902 GTGCTCTCTGTCACCTGACTCTTCTGAAGGCAACTAATCTTAAGCGTAAAGTTTATCA 3961
Qy 1626 CTCAGTCAGTCGACAGCATTTAG-AAAAAATTTCAAAAGTCTGCCCGTAGCCCGGAGCAAA 1684
Db 3962 CTCAGTCAGTCGACAGCATTTAGAAAAAATTTCAAAAGTCTGCCCGTAGCCCGGAGCAAA 4021
Qy 1685 ACTTAGAAGCCCTATTGAACCTTGGCAACVTCGGTTTTTATATAGAGATCAGGAGGAGC 1744
Db 4022 ACTTAGAAGCCCTATTGAACCTTGGCAACVTCGGTTTTTATATAGAGATCAGGAGGAGC 4081
Qy 1745 AGGCGGAACAGGACAAAACGGGATTAATAAAGGACCGCTTTTAGTCAATGACCCCTCAGG 1804
Db 4082 AGGCGGAACAGGACAAAACGGGATTAATAAAGGACCGCTTTTAGTCAATGACCCCTCAGG 4141
Qy 1805 CAAATGGA-CTTTGGAGGCTCTGGAAAAAGGAAAAAGCTGGGCAAAATGAATGCTTAATAGG 1864
Db 4142 CAAATGGA-CTTTGGAGGCTCTGGAAAAAGGAAAAAGCTGGGCAAAATGAATGCTTAATAGG 4201
Qy 1865 GCTTCTTCCAGTGGGTCATCAAGGACACTTTTAAAAAAGATTTGTCCAAGTGAAGTAAG 1924
```

Db 4202 GCTTGTCTCAGTGGGCTCTACAAGGACACTTTTAAAAAGATTGTCCNAGTAGAGTAAG 4261  
QY 1925 CCGGCCCTTCGTCTCATGCCCCCTTATTTCAGAGGAATCACTGGAAGGCCACTGCCCCAGG 1984  
Db 4262 CCGGCCCTTCGTCTCATGCCCCCTTATTTCAGAGGAATCACTGGAAGGCCACTGCCCCAGG 4321  
QY 1985 GGACAAAGGTCCTTGTGATCAGAGGACCACTAACAGATGATCCAGCAGCAGGACTGAGGG 2044  
Db 4322 GGACAAAGGTCCTTGTGATCAGAGGACCACTAACAGATGATCCAGCAGCAGGACTGAGGG 4381  
QY 2045 TGCTGGGGCAAGCGCCATCCCATGCCATCACCTCTACAGAGCCCTGGGTATGCTTGACC 2104  
Db 4382 TGCTGGGGCAAGCGCCATCCCATGCCATCACCTCTACAGAGCCCTGGGTATGCTTGACC 4441  
QY 2105 ATTGAGGGCCAGGAAGGTTGCTCTCTGGAACACTTGCTGGCGTCTTCTTACTCTTTC 2164  
Db 4442 ATTGAGGGCCAGG-AGGTTGCTCTCTGGAACACTTGCTGGCGTCTTCTTACTCTTTC 4500  
QY 2165 TGTCCCGGCAACTGTCTCCAGATCTGTCACATTAATCTTGTAGGGGGTCCNTAAGACGGCA 2224  
Db 4501 TGTCCCGGCAACTGTCTCCAGATCTGTCACATA-TCTGAGGGGGTCC-TAAGACGGGCA 4558  
QY 2225 GTCACATAGATPACTTTTCCAGGCACTAACTTATGAACCTGGGAGCTTTTATTCTTTTCAC 2284  
Db 4559 GTCACATAGATAC-TTCTCCAGGCACTAAGTTATG-ACTGGGAGCTTTTATTCTTTTCAC 4616  
QY 2285 ATGCTTTTCTPAATATGCTTTGAAAGCCCACTACCTTGTGTAGGAGAGACATTTCTAGCAA 2344  
Db 4617 ATGCTTTTCTPAATATGCTTTGAAAGCCCACTACCTTGTGTAGGAGAGACATTTCTAGCAA 4676  
QY 2345 AAGCAGGGCCATATACACCTGAACATAGGAGAGAACACCCGTTTGTGTCNCCCTG 2404  
Db 4677 AAGCAGGGCCATATACACCTGAACATAGGAGAGAACACCCGTTTGTGTTGT-CCCTG 4735  
QY 2405 CTTGAGGAAGGAATTAATCTGAAAGTCTGGGCAACAGAGGACAAATATGACAGGCCAAA 2464  
Db 4736 CTTGAGGAAGGAATTAATCTGAAAGTCTGGGCAACAGAGGACAAATATGACAGCG-CAAA 4794  
QY 2465 GAATGCCCGTCTGTTTCAAGTTTAACTTAAAGGATTCACCTTCTTCCCTTACCAAAAGGCA 2524  
Db 4795 GAATGCCCGTCTGTTTCAAGTTTAACTTAAAGGATTCACCTTCTTCCCTTACCAAAAGGCA 4854  
QY 2525 GTACCCCTCTCAGACCCCAAGGCCCAACAGAGATTCCAAAAGATTGTTAAGGACTTAAAGC 2584  
Db 4855 GTACCCCTCTCAGACCCCAAGGCCCAACAGGACTTCCAAAAGATTGTTAAGGACTTAAAGC 4914  
QY 2585 CCAAGGCTTAGTAAAAACCATGACATACTCCCTGACGTAATTCGCTAGTGGATTGAGGAG 2644  
Db 4915 CCAAGGCTTAGTAAAAACCATGACATACTCCCTGACGTAATTCGCTAGTGGATTGAGGAG 4967  
QY 2645 CACAGAAACCCAGTGACAGTGGAGGTTAGTGCAAGATCTCAGGATTAATCAATGGAGGC 2704  
Db 4968 TACAGAAACCCACAGACAGTGGA-GGTTAGTGCAAGATCTCAGGATTAATCAAT-AGGCG 5025  
QY 2705 CGTTGTCTTTTATACCCAGCTGTACTAGCCCTTATACGTGMYTTCACAAATACAGA 2764  
Db 5026 TGTGTTCTCTATAGCAGCTGTACTAGCCCTTATCTCTGTTTCCCAATATACAGA 5085  
QY 2765 GGAAGCAGAGTGGTTTACASTCTCTGACCTTMAAGGATGCTTCTCTGCAATCCCTGTACA 2824  
Db 5086 GGAAGCAGAGTGGTTTACAGTCTCTGACCTTTCAGGATGCTTCTCTGCAATCCCTGTACA 5145  
QY 2825 TCCTGACTCTCAATTTCTTGTGTTGCCCTTTGAAGATATCTTCAAAACCCACATCTCAACTCAC 2884  
Db 5146 TCCTGACTCTCAATTTCTTGTGTTGCCCTTTGAAGATATCTTCAAAACCCACATCTCAACTCAC 5205  
QY 2885 CTGCACTRTTTTACCCCAAGGTTTCAGGGATAGYCCCCATCTATTATTTGGCCAGGCATTAGC 2944  
Db 5206 CTGCACTATTATACCCCAAGGTTTCAGGGATAGTCCCCATCTATTATTTGGCCAGGCATTAGC 5265  
QY 2945 CCAAGACTTGAGYCATYMTCATACCTGGACACTCTTGCTCTTCRTGTAGTGATGATTT 3004

Db 5266 CCAAGACTTGAGGCCAATCTCATACCTGGACA--CTTGTCTTTCGGTAGTGATGATTT 5323  
QY 3005 ACTTTTGGCYGCCRRTTCAGAAAACCTTTGTGCCATCAAGCCACCACCAAGCRCTCTTTAAATTT 3064  
Db 5324 ACTTTTGGCGGCCATTTCAGAAAACCTTTGTGCCATCAAGCCACCACCAAGCGCTCTTCAATTT 5383  
QY 3065 CCTCGCYACTGTGGCTACAWGGTTTCCAAAASABARGCTCARTCTGTCTCAGACGAGGT 3124  
Db 5384 CCTCGCTACTGTGGCTACATGGTTTCCAAAACCAAAAGGCTCAACTCTGTCTCAGCAGGTT 5443  
QY 3125 TAAATACTTAGGCTTAARATTATCCAAAGCACCARGGCCCTCAGTGAGGAAVRYATCCA 3184  
Db 5444 ----TACTTAGGCTTAARATTATCCAAAGCACCARGGCCCTCAGTGAGGAAACATATCCA 5499  
QY 3185 GCCTATCTAGGCTTATCTCATCTCYCAAAACCCCTTAAAGCAACTAAGRRRTTCTTGGCCT 3244  
Db 5500 GCCTATCTAGGCTTATCTCTCATCCCAAAACCCCTTAAAGCAACTAAGGGGATTCCTTGGCGT 5559  
QY 3245 AAVAGGTYTTCGCCGAATGATTCGCCAGGTTCGCCAGGTTCGCCAATATAGCCAGGYCATTAWATA 3304  
Db 5560 AATAGGTTTTCGCCGAATGATTCGCCAGGTTCGCCAGGTTCGCCAATATAGCCAGGTCAATTAATA 5618  
QY 3305 CASTAATTAAGGAACTCAGAAAGCCCAATACCCATTTARTAAGATGGAAMCTGAAGYMR 3364  
Db 5619 CACTAATTAAGGAACTCAGAAAGCCCAATACCCATTTAGTAGATGGAACACTGAAGTAG 5678  
QY 3365 AAGTGGCTTTTCAGGCCCCCTTAAAGAGGCTTTAAACCCCAAGYCCCAAGTGTAAAGYTTGCC 3424  
Db 5679 AAGTGGCTTTTCAGGCCCC-----TAAACCCCAAGGCCAGTGTAAAGTTTGCC 5724  
QY 3425 AACRGCGCAAGACTTTTSTTYATATYRTACAGAAAAAACAAGAAAYAGCTCTRGGAGTCC 3484  
Db 5725 AACAGGGCAAGACTTTTCTTCATATGTACAG-AAAAAACAAGGAATAGCTCTAGGAGTCC 5783  
QY 3485 TTACACAGTCCRAGGAYGAGCTTCAACCCYRTGCRVACCTGASTAAGGAAAYTGATG 3544  
Db 5784 TTACACAGATCCGAGGGATGAGCTTGCAACCTGTGGCATACCTGACTAGGAAATTTGATG 5843  
QY 3545 TAGTGGCAAGGTTTGRCTCATTTGTTAYGGGTAGTGGGAGTAGCAGTGYKTATAGTAT 3604  
Db 5844 TAGTGGCAAGGTTTGACCTCATTTGTTACGGGTAGTGGTGGCAGTAGCAGTCTTAGTAT 5903  
QY 3605 CTGAAGCAGTTAAATAATATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAGTGA 3664  
Db 5904 CTGAAGCAGTTAAATAATATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGAGTGA 5963  
QY 3665 AYRGCATACTCATCTAAGGAGACTTGTGCTGTGACACACACACACACACACACACACACAC 3724  
Db 5964 ATGGCATACTCATCTGTAAGGAGACTTGTGGCTGTGACACACACACACACACACACACAC 6023  
QY 3725 AGGCTCTATTACTTGAARGGCCAGTGTGCTGCACTGTGCACTTGTGCAACTCTTTAAACCCAG 3784  
Db 6024 AGGCTCTATTACTTGAARGGCCAGTGTGCTGCACTGTGCACTTGTGCACTCTTTAAACCCAG 6083  
QY 3785 YNCATTTCTTCAGACAAATGAAGAAAAAGATAPAAAYATAACTGTCTCAACAAATTAATTTCTC 3844  
Db 6084 CCACATTTCTTCAGACAAATGAAGAAAAAGATAPAAACATAAATGTCTCAACAAAGTAATTTCTC 6143  
QY 3845 AAACCTATGCCACTCGAGGGACCTTGTAGABRTTCCYTTGACTGATCCYAGACCTTCAAC 3904  
Db 6144 AAACCTATGCCACTCGAGGGACCTTTTAGAGGTTCTTTGACTGATCCCGACC-TCAAC 6202  
QY 3905 TTGTATACTGATGGAAGTTCTTTTGTAGAAAAAGGACTTCGAAAAAGYGGGGTATGAGTGG 3964  
Db 6203 TTGTATACTGATGGAAGTTCTTTTGTAGAAAAAGGACTTCGAAAAAGYGGGGTATGAGTGG 6262  
QY 3965 GTCAAGTAAATGGAATATTTGAAAGTAATCCCTCCTCCTCAGGAACTAGTCTYAGCTTR 4024  
Db 6263 GTCAAGTAAATGGAATATTTGAAAGTAATCCCTCCTCCTCAGGAACTAGTCTYAGCTTR 6322  
QY 4025 GCAGAACTAATAGCCYTCATYTKGGGCACTAGAAATTTAGGAGAAAGRAAAAGGGYAAATATA 4084  
Db 6323 GCAGAACTAATAGCCCTCATCTTGGGCACTAGAAATTTAGGAGAAAGRAAAAGGGYAAATATA 6382







[illegible]

Db	9621	ACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATTCAGAGAGCTC	9688
Qy	7384	ACTAAATGCTAAATTAGGCAAGACGAGAGTAAAGAAATAGCCAATCATYATATGCMTG	7443
Db	9681	ACTAAATGCTAATTAGGCAAGACGAGAGTAAAGAAATAGCCAATCATYATATGCMTG	9740
Qy	7444	AGACACAGCAGGAGGAGCAATGATCGGATATAAAACCAAGTTCGAGCCCGCAACGG	7503
Db	9741	AGACACAGCAGGAGGAGCAATGATCGGATATAAAACCAAGTTCGAGCCCGCAACGG	9800
Qy	7504	CAACCCCTTTGGTCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCT	7563
Db	9801	CAACCCCTTTGGTCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACTCT	9860
Qy	7564	ATTAATCTTGCARCTGCR	7582
Db	9861	ATTAATCTTGCARCTGCA	9879
RESULT 13			
LOCUS	AR699757	10499 bp	DNA linear PAT 14-SEP-2005
DEFINITION	Sequence 3 from patent US 6919438.		
ACCESSION	AR699757		
VERSION	AR699757.1	GI:75205546	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 10499)		
AUTHORS	Alliel,P.M., Perin,J.-P. and Rieger,F.		
TITLE	Nucleic sequence and deduced protein sequence family with human		
JOURNAL	endogenous retroviral motifs, and their uses		
	Patent: US 6919438-A 3 19-JUL-2005;		
	Institut National de la Sante et de la Recherche Medical-Inserm;		
	Paris;		
	FRX;		
FEATURES	Location/Qualifiers		
source	1..10499		
	/organism="unknown"		
ORIGIN	/mol_type="genomic DNA"		
Query Match	85.9%;	Score 6436;	DB 2; Length 10499;
Best Local Similarity	95.7%;	Pred. No. 0;	
Matches 6680;	Conservative 195;	Mismatches 50;	Indels 54; Gaps 28;
Qy	606	GATGGGAAACGTTCCCGCAAGACAAACCGCCCTTAAGACGTAATTCGAGAAATTTGGGA	665
Db	2953	GATGGGAAACGTTCCCGCAAGACAAACCGCCCTTAAGACGTAATTCGAGAAATTTGGGA	3012
Qy	666	MCAATTTGACCCCTCAGACACTAGAAAGAAACGACTTATATCTTCGACGTCGCCCTTG	725
Db	3013	CCAAATTTGACCCCTCAGACACTAGAAAGAAACGACTTATATCTTCGACGTCGCCCTTG	3072
Qy	726	GCATCTCTGAGGAAAGTATAAAATATAACCAATCTTACAGCTAGACYCTTTTGTAGAA	785
Db	3073	GCATCTCTGAGGAAAGTATAAAATATAACCAATCTTACAGCTAGACCTCTTTTGTAGAA	3132
Qy	786	AAGCAAAATGGAGTGCCTAAGTACAACTTTCTTTTTCATTAAGAGACAACTCAC	845
Db	3133	AAGCAAAATGGAGTGCCTAAGTACAACTTTCTTTTTCATTAAGAGACAACTCAC	3192
Qy	846	AATTATGTAAGAAAGTGTATTATGCGCTTACAGGAAGCCTTCAGAGTCTACCTCCCTATC	905
Db	3193	AATTATGTAAGAAAGTGTATTATGCGCTTACAGGAAGCCTTCAGAGTCTACCTCCCTATC	3252
Qy	906	CCAGCATCCCGACTCCTTCCCCCACTAATAGGACCCCTTCAACCCCAATGGTCCAA	965
Db	3253	CCAGCATCCCGACTCCTTCCCCCACTAATAGGACCCCTTCAACCCCAATGGTCCAA	3312
Qy	966	AAGGAGATAGACAAAAGGGGTAAACAGTGAACCAAGAGTGCCAAATTTCCCAATTTATGA	1025

Db 3313 AAGGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTGCCTAATATTCCTCCCAATTATGA 3372  
Qy 1026 CCCCTCCCAAGCAGTGGGAGGAGAGAAATTCGGCCACGACGAGTGCATGCTGCTTTTYY 1085  
Db 3373 CCCCT-CCAAGCAGTGGGAGGAGAGAAATTCGGCCACGACGAGTGCATGCTGCTTTTTC 3431  
Qy 1086 TCTCCAGACTTAAAGCAAAATAAAACAGACTTACGTAAATCTCAGATAAATTCCTGATGG 1145  
Db 3432 TCTCCAGACTTAAAGCAAAATAAAACAGACTTACGTAAATCTCAGATAAATTCCTGATGG 3491  
Qy 1146 CTATATGTTGTTTACAGGGTTAGGCAATTCCTGATCTGACATGAGAGATATATA 1205  
Db 3492 CTATATGATGTTTACAGGGTTAGGCAATTCCTGATCTGACATGAGAGATATA-A 3550  
Qy 1206 TGTCACTGCTAAATCAGACATAAACCCTAATGAGAGAGTGCACCACTAATCTGAGCCT 1265  
Db 3551 TGTCACTGCTAAATCAGACATAAACCCTAATGAGAGAGTGCACCACTAATCTGAGCCT 3610  
Qy 1266 GAGGTGTTGGCGATCTCTGGTATCTCAGTCAGTCAATGATGATGATGATGATGATGATG 1325  
Db 3611 GAGGTGTTGGCGATCTCTGGTATCTCAGTCAGTCAATGATGATGATGATGATGATGATG 3667  
Qy 1326 AAGAGAAATGATTCCTCCACAGGCCAGCAGTTCCTCCAGTCTASACCTCAATTCGGGAC 1385  
Db 3668 AAGAGAAATGATTCCTCCACAGGCCAGCAGTTCCTCCAGTCTAGACCTCAATTCGG- 3724  
Qy 1386 ACAGAAATCAGTAACATGAGGAGATGTTGCTGTCAGACATTTGCTAACTTGTGCTASAA 1445  
Db 3725 ACACAGAATCAGAAAT-GGAGATTTGTTGCTGTCAGACATTTGCTAACTTGTGCTAGAA 3783  
Qy 1446 GGACTAAGGAAACTASGAGAAATCTAYGAATTAATCAATGATGTCACCATCAACACA 1505  
Db 3784 GGACTAAGGAAACTASGAGAAATCTAYGAATTAATCAATGATGTCACCATCAACACA 3842  
Qy 1506 GGGGAGGGAAGAAATCTACTGCTTTCTGGAGAGACTAAGGAGGAGTTCAGGAGGAGC 1565  
Db 3843 -GGGAGGGAAGAAATCTACTGCTTTCTGGAGAGACTAAGGAGGAGTTCAGGAGGAGC 3901  
Qy 1566 GTGCTCTCTGTCACTGACTCTCTGAAGGCAACTAATCTTAAGCGTAAGTTTATCA 1625  
Db 3902 GTGCTCTCTGTCACTGACTCTCTGAAGGCAACTAATCTTAAGCGTAAGTTTATCA 3961  
Qy 1626 CTCAGTCAGTCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGAGGAGGAGC 1684  
Db 3962 CTCAGTCAGTCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGAGGAGGAGC 4021  
Qy 1685 ACTTAGAAACCTTATTTGAACCTTGGCAACCTGCTGCTTTTATATAGAGATCAGGAGGAGC 1744  
Db 4022 ACTTAGAAACCTTATTTGAACCTTGGCAACCTGCTGCTTTTATATAGAGATCAGGAGGAGC 4081  
Qy 1745 AGGCGGAACAGGACAAACCGGATTAATAAAAGGCGCCGCTTTATAGTCAAGCCCTCAGG 1804  
Db 4082 AGGCGGAACAGGACAAACCGGATTAATAAAAGGCGCCGCTTTATAGTCAAGCCCTCAGG 4141  
Qy 1805 CAAGTGGACTTTGGAGGCTCTGGAAGGGAAGCTGGGCAATTTGATGCTAATAGG 1864  
Db 4142 CAAGTGGACTTTGGAGGCTCTGGAAGGGAAGCTGGGCAATTTGATGCTAATAGG 4201  
Qy 1865 GCTTGTCTCCAGTGGGCTTACAGGACACTTTAAAAAGATTTGTCGAAGTAAAGTAAAG 1924  
Db 4202 GCTTGTCTCCAGTGGGCTTACAGGACACTTTAAAAAGATTTGTCGAAGTAAAGTAAAG 4261  
Qy 1925 CCGCCCTTGTGCTAATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGG 1984  
Db 4262 CCGCCCTTGTGCTAATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGG 4321  
Qy 1985 GGACAAAGTCTTTTGAAGTCAAGGACACTTAACAGATGATCCAGCAGGACTTGAGG 2044  
Db 4322 GGACAAAGTCTTTTGAAGTCAAGGACACTTAACAGATGATCCAGCAGGACTTGAGG 4381  
Qy 2045 TGCCTGGGCAAGCGCATCCCATGCTATCACCCTCAGAGGCCCTGGGTATGCTTGACC 2104  
Db 4382 TGCCTGGGCAAGCGCATCCCATGCTATCACCCTCAGAGGCCCTGGGTATGCTTGACC 4441

Qy 2105 ATTGAGGCCAGGAGGTTGCTCTCTGACACTGCTGGGTCTTCTTAGTCTTACTCTTTC 2164  
Db 4442 ATTGAGGCCAGG-AGGTTGCTCTCTGACACTGCTGGGTCTTCTTAGTCTTACTCTTTC 4500  
Qy 2165 TGTCCCGACAACTGTCTCTCAGACTCTCTCACTATTCTTGGGGGCTCCTTAAGAGGGCA 2224  
Db 4501 TGTCCCGACAACTGTCTCTCAGACTCTCTCACTA-TCTTGGGGGTCC-TAAGACGGCA 4558  
Qy 2225 GTCACTAGATCTTTTCTCCAGCCACTAAGTTATGAATGAACTGGGGAGCTTTATCTTTTCA 2284  
Db 4559 GTCACTAGATAC-TTCTCCAGCCACTAAGTTATG-ACTTGGGGAGCTTTATCTTTTCA 4616  
Qy 2285 ATGCTTTTCTAATATATGCTTGAAGCCCACTACCTTGTAGGGAGAGACATTTCTAGCAA 2344  
Db 4617 ATGCTTTTCTAATATATGCTTGAAGCCCACTACCTTGTAGGGAGAGACATTTCTAGCAA 4676  
Qy 2345 AAGCAGGCCCAATATACACCTGAACATAGAGAGAGAAACACCGTTTGTGTTNCCCTG 2404  
Db 4677 AAGCAGGCCCAATATACACCTGAACATAGAGAGAGAAACACCGTTTGTGTT-CCCCTG 4735  
Qy 2405 CTGAGGAGGAAATTAATCTGAACTTGGGCAACAGAGGACAATATGGAAGAGCAAA 2464  
Db 4736 CTGAGGAGGAAATTAATCTGAACTTGGGCAACAGAGGACAATATGGAAGAG-CAAA 4794  
Qy 2465 GAATGCCCTCTGTTCAAGTTAACTAAAGATTTCCACTTCTTCTTCCCTACCAAGGCA 2524  
Db 4795 GAATGCCCTCTGTTCAAGTTAACTAAAGATTTCCACTTCTTCTTCCCTACCAAGGCA 4854  
Qy 2525 GTACCCCTCAGACCCCAAGGCCCAACAGGATTTCCAAAGATTTGTTAAGGACTTAAAGC 2584  
Db 4855 GTACCCCTCAGACCCCAAGGCCCAACAGGACTTCCAAAGATTTGTTAAGGACTTAAAGC 4914  
Qy 2585 CCAAGGCTTAGTAAACCAATGATTAATCTCCGTGAGTAAATTCCTGATGAGTGGAGG 2644  
Db 4915 CCAAGGCTTAGTAAACCAATGATTAATCTCCGTGAGTAAATTCCTGATGAGTGGAGG 4967  
Qy 2645 CACAGAAACCCAGTGAGGAGTGTAGTGAAGATCTCAGAGTATATCAATGAGGAGC 2704  
Db 4968 TACAGAAACCCAGTGAGGAGTGTAGTGAAGATCTCAGAGTATATCAAT-GAGGC 5025  
Qy 2705 GGTGTCTTTTATACCCAGCTGTACTAGCCCTTATATCTGTGTTTCCCAATACCA 2764  
Db 5026 TGTGTTCTCTATAGCCAGCTGTACTAGCCCTTATATCTGTCTTCCCAATACCA 5085  
Qy 2765 GGAAGCAGTGGTTTACCTCTGGAACCTTGAAGATGCTTCTTCTGATCCCTGTACA 2824  
Db 5086 GGAAGCAGTGGTTTACAGTCTCTGGAACCTTGAAGATGCTTCTTCTGATCCCTGTACA 5145  
Qy 2825 TCCTGACTCTCAATTTCTTGTGCTTGAAGATACTTCAACCCAGTCACTCACTCAC 2884  
Db 5146 TCCTGACTCTCAATTTCTTGTGCTTGAAGATACTTCAACCCAGTCACTCACTCAC 5205  
Qy 2885 CTGGACTTTTACCCCAAGGTTTCAGGATAGYCCCATCTATTTGGCAGGACTTAGC 2944  
Db 5206 CTGGACTTTTACCCCAAGGTTTCAGGATAGTCCCATCTATTTGGCAGGACTTAGC 5265  
Qy 2945 CCAAGACTTGAGCCAACTCTCATCTCTGAGCA-CTTGTCTTCTCGTAGGTGGATTT 3004  
Db 5266 CCAAGACTTGAGCCAACTCTCATCTCTGAGCA-CTTGTCTTCTCGTAGGTGGATTT 5323  
Qy 3005 ACTTTTTCGCTTTCAGAAACCTTGTGCTTCAAGCCAGGAGTCTCTTAAATTT 3064  
Db 5324 ACTTTTTCGCTTTCAGAAACCTTGTGCTTCAAGCCAGGAGTCTCTTAAATTT 5383  
Qy 3065 CCTCCYACCTGTGGCTACAWGTTTCCAAACAPARGCTCACTCTGTCTCAGCAGAGT 3124  
Db 5384 CCTCCYACCTGTGGCTACATGTTTCCAAACAGGCTCACTCTGTCTCAGCAGAGT 5443  
Qy 3125 TAAATACCTTAGGCTTAATATCAAGGCAACAGGCCCCCTCAGTGAAGAAATATCA 3184  
Db 5444 ----TACTTAGGCTTAAATATCAAGGCAACAGGCCCCCTCAGTGAAGAAATATCA 5499



Db	7643	CATCCCTAGATACATCTCGGGAAGGACCCCTACCGAGTCATTTTATCTACCCCAACTGCGG	7702
Qy	5404	TTAAAGTGGCTGGAGTGGAGTCTTTGGATACATCACACTTGAGTCAAAATCCTGGATACTGC	5463
Db	7703	TTAAAGTGGCTGGAGTGGAGTCTTTGGATACATCACACTTGAGTCAAAATCCTGGATACTGC	7762
Qy	5464	CAAAAGGAACCTGAAAATCCAGGAGACAAACGCTAGCTATTCTCTGTGAACCTCTAGAGGATT	5523
Db	7763	CAAAAGGAACCTGAAAATCCAGGAGACAAACGCTAGCTATTCTCTGTGAACCTCTAGAGGATT	7822
Qy	5524	TGCGCTGCTCTTTCAACAACAACCCAGGAGGAAGTAACATAAATCATAAATCCCCCATG	5583
Db	7823	TGCGCTGCTCTTTCAACAACAACCCAGGAGGAAGTAACATAAATCATAAATCCCCCATG	7881
Qy	5584	GSCTCCCTTATCATATTTTTCTCTASTGTTSTTTACCTSTTTCACTCTCACTGCA	5643
Db	7882	GSCTCCCTTATCATATTTTTCTCTASTGTTSTTTACCTSTTTCACTCTCACTGCA	7941
Qy	5644	CCCCCTCCATGCGCGCTGATGACGAGTAGCTCCCTYACCMAGAGTTTCTATGGAAGATTG	5703
Db	7942	CCCCCTCCATGCGCGCTGATGACGAGTAGCTCCCTYACCMAGAGTTTCTATGGAAGATTG	8001
Qy	5704	CAGGCTCCGGAATATTTGATGCCCCCATCGTATAGGAGTCTTTSTAAGGAAACCCCAACC	5763
Db	8002	CAGGCTCCGGAATATTTGATGCCCCCATCGTATAGGAGTCTTTSTAAGGAAACCCCAACC	8061
Qy	5764	TTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCAAT	5823
Db	8062	TTCACTGCCCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCAAT	8121
Qy	5824	GCAAAATCACTAATTTGGACAGGAAAAAATGATTAATCCTTAGTTGTCTCGGAGGACTTGA	5883
Db	8122	GCAAAATCACTAATTTGGACAGGAAAAAATGATTAATCCTTAGTTGTCTCGGAGGACTTGA	8181
Qy	5884	GTCACTGTCTGTTGACTTACTTCAACCCAACTGGTATGCTGATGCGGGGTGAGATTCAA	5943
Db	8182	GTCACTGTCTGTTGACTTACTTCAACCCAACTGGTATGCTGATGCGGGGTGAGATTCAA	8241
Qy	5944	GATCAGGCACAGAAAAACATGTAAGAAGTAGTATCTCCCAACTCACCCGGGTACATGCC	6003
Db	8242	GATCAGGCACAGAAAAACATGTAAGAAGTAGTATCTCCCAACTCACCCGGGTACATGCC	8301
Qy	6004	ACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAAATACATGAAAAACCTCCGTAACCCAT	6063
Db	8302	ACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAAATACATGAAAAACCTCCGTAACCCAT	8361
Qy	6064	ACTCGCTGTGTAAGCCTATTTAATPACCAACCTCTCACTGGGTCCAATGAGTCTCGGGCCAA	6123
Db	8362	ACTCGCTGTGTAAGCCTATTTAATPACCAACCTCTCACTGGGTCCAATGAGTCTCGGGCCAA	8421
Qy	6124	AACCTACTAAGTGTGGATATGCTCCCTGAACTTCARGCCATATGTTTCAATCCCT	6183
Db	8422	AACCTACTAAGTGTGGATATGCTCCCTGAACTTCARGCCATATGTTTCAATCCCT	8481
Qy	6184	GTACTGAAACAATGGAAACAACTTCAGCACAGAAAAAACAACCACTTCCTGTTTAGTAGGA	6243
Db	8482	GTACTGAAACAATGGAAACAACTTCAGCACAGAAAAAACAACCACTTCCTGTTTAGTAGGA	8541
Qy	6244	CCTCTGTTTCCAAATSTGGAATAACCCATACCTCAAAACCTCACCTGTGTAAATTTTAC	6303
Db	8542	CCTCTGTTTCCAAATSTGGAATAACCCATACCTCAAAACCTCACCTGTGTAAATTTTAC	8601
Qy	6304	AATACTACATACACAAACCACTCCCAATGCATGAGGTGGGTAACTCTCCCAACAATA	6363
Db	8602	AATACTACATACACAAACCACTCCCAATGCATGAGGTGGGTAACTCTCCCAACAATA	8661
Qy	6364	GTCTGCTTACCTCAGGAATATTTTTGTCTGTGTACTCAGGCTATCGTTGTTGAAAT	6423
Db	8662	GTCTGCTTACCTCAGGAATATTTTTGTCTGTGTACTCAGGCTATCGTTGTTGAAAT	8721
Qy	6424	GGCTCTTCAGAAATCTATGTGCTCTCTCATTTCTTAGTGCCCCCYATGRCATCTACACT	6483
Db	8722	GGCTCTTCAGAAATCTATGTGCTCTCTCATTTCTTAGTGCCCCCYATGRCATCTACACT	8781

Qy	6484	GAACAAGATTTATACAGTTAATGTCATATCTAAAGCCCGCGCAACAAAAGAGTACCCATTCCTT	6543
Db	8782	GAACAAGATTTATACAGTTAATGTCATATCTAAAGCCCGCGCAACAAAAGAGTACCCATTCCTT	8841
Qy	6544	CCTTTTGTTATAGGAGCAGAGTGCTAGGTGCACTAGGTACTGGCAATTTGGCGGTATACACA	6603
Db	8842	CCTTTTGTTATAGGAGCAGAGTGCTAGGTGCACTAGGTACTGGCAATTTGGCGGTATACACA	8901
Qy	6604	ACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGACACTGGAAACGGGTC	6663
Db	8902	ACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGACACTGGAAACGGGTC	8961
Qy	6664	GCGACTCCCTGGTGCACCTTTGCAAGATCAACTTAACTCCTTAGCAGCAGTAGTCCTTCRA	6723
Db	8962	GCGACTCCCTGGTGCACCTTTGCAAGATCAACTTAACTCCTTAGCAGCAGTAGTCCTTCAA	9021
Qy	6724	AATCGAAGAGCTTTAGACTVGCCTAAACCGCTGABAGAGGGGGAACCTGTTTTATTTTTAGGG	6783
Db	9022	AATCGAAGAGCTTTAGACTVGCCTAAACCGCTGABAGAGGGGGAACCTGTTTTATTTTTAGGG	9081
Qy	6784	GAAGAATGCTGTTATTTATGTTTAACTCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT	6843
Db	9082	GAAGAATGCTGTTATTTATGTTTAACTCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT	9141
Qy	6844	CSAGATCGAATACAAACGTAKAGCAGAGAGCTTCGAAACAACCTGAGACCTCGGGGCTCCTC	6903
Db	9142	CGAGATCGAATACAAACGTAGAGCAGAGAGCTTCGAAACAACCTGAGACCTCGGGGCTCCTC	9201
Qy	6904	AGCCRAATGGATGCCCTGGATTCTCCCTTCTTAGGAGCTCTAGGAGCTATAATATTCCTA	6963
Db	9202	AGCCRAATGGATGCCCTGGATTCTCCCTTCTTAGGAGCTCTAGGAGCTATAATATTCCTA	9261
Qy	6964	CTCCTCTTTTGGACCCCTGTATCTTTTTRACCTCCTTGTGTTAACTTGTCTCTCCAGAACTCGAA	7023
Db	9262	CTCCTCTTTTGGACCCCTGTATCTTTTAACTCCTTGTGTTAACTTGTCTCTCCAGAACTCGAA	9321
Qy	7024	GCTGTAAACTTACAAATGGAGCCCAAGATGCAGTCCAAAGACTAAGATCTACCGCAGACCC	7083
Db	9322	GCTGTAAACTTACAAATGGAGCCCAAGATGCAGTCCAAAGACTAAGATCTACCGCAGACCC	9381
Qy	7084	CTGGACCGGCTGYTAGCCCAACGATCTGATGTTTAATGACATCAAAAGCAACCCCTCTCTGAG	7143
Db	9382	CTGGACCGGCTGYTAGCCCAACGATCTGATGTTTAATGACATCAAAAGCAACCCCTCTCTGAG	9441
Qy	7144	GAATCTCAGCTGCACAACTCTACTAGGCCCCCAATTCAGCAGGAAACGAGTTAGAGCGGT	7203
Db	9442	GAATCTCAGCTGCACAACTCTACTAGGCCCCCAATTCAGCAGGAAACGAGTTAGAGCGGT	9501
Qy	7204	SGTCGGCCAACTCCCCAAACAGCACTTAGGTTTTCCGTTGTGAGATGGGGACTGAGAGAC	7263
Db	9502	C-TGGGCCAACTCCCCAAACAGCACTTAGGTTTTCCGTTGTGAGATGGGGACTGAGAGAC	9560
Qy	7264	AGGACTAGCTGGAATTTCTTAGGCTGAYTAAGAAATCCYTAAGCCTTAGSTGGGAAGGTGACC	7323
Db	9561	AGGACTAGCTGGAATTTCTTAGGCTGAYTAAGAAATCCYTAAGCCTTAGSTGGGAAGGTGACC	9620
Qy	7324	ACATCCACTTTTAAACACGGGGCTTGCAACTTAGYTACACCTGACCAATCAGAGAGCTC	7383
Db	9621	ACATCCACTTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTC	9680
Qy	7384	ACTAAATGCTAAATTTAGGCAAGACAGGAGTAAAGAAATAGCCAAATCATVTATTTGCGTG	7443
Db	9681	ACTAAATGCTAAATTTAGGCAAGACAGGAGTAAAGAAATAGCCAAATCATVTATTTGCGTG	9740
Qy	7444	AGAGCACAGCAGGAGGACAAATGATCGGGATATAAAACCCAGTCTTCGAGCCGCAACGG	7503
Db	9741	AGAGCACAGCAGGAGGACAAATGATCGGGATATAAAACCCAGTCTTCGAGCCGCAACGG	9800
Qy	7504	CAACCCCTTTGGGTCCCTCTCCCTTTGTATGGGAGCTCTGTTTTTCATGCTATTTCACTCT	7563
Db	9801	CAACCCCTTTGGGTCCCTCTCCCTTTGTATGGGAGCTCTGTTTTTCATGCTATTTCACTCT	9860

```
QY 7564 ATTAATCTTGCARCTGCR 7582
|||||:|||||:|||||:
Db 9861 ATTAATCTTGCARCTGCA 9879

RESULT 14
AX007980 10499 bp DNA linear PAT 06-SEP-2000
LOCUS Sequence 3 from Patent WO967395.
DEFINITION AX007980
ACCESSION AX007980.1 GI:9995677
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1
REFERENCE
AUTHORS Perin, J.P., Rieger, F. and Alliel, P.M.
TITLE Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses
JOURNAL Patent: WO 9967395-A 3 29-DEC-1999;
INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER
FRANCOIS (FR); ALLIEL PATRICK M (FR)
FEATURES
source
1..10499
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 85.9%; Score 6436; DB 2; Length 10499;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 6680; Conservative 195; Mismatches 50; Indels 54; Gaps 28;

QY 606 GATGGGAACGTTCCCGCAAGACAAACGCGCCCTAAGACGATTTCTGCGARAAATGGGA 665
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2953 GATGGGNAAGTTCCCGCAAGACAAACGCGCCCTAAGACGATTTCTGCGARAAATGGGA 3012
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 666 MCAATTTGACCTCAGACACTAAGAAAGAACGACTTATATTTCTGCGAGTGCCTGCTG 725
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3013 CCAATTTGACCTCAGACACTAAGAAAGAACGACTTATATTTCTGCGAGTGCCTGCTG 3072
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 726 GCACTCTGAGGAAGTATAAATTATAACACCATTTACAGTAGACATCTTTGTTAGAA 785
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3073 GCATCTCTGAGGAAGTATAAATTATAACACCATTTACAGTAGACATCTTTGTTAGAA 3132
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 786 AAGCAAAATGGAGTGAAGTGCCTAAGTACAAATCTTTTCAATTAAGAGACAATCTAC 845
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3133 AAGCAAAATGGAGTGAAGTGCCTAAGTACAAATCTTTTCAATTAAGAGACAATCTAC 3192
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 846 AATTATGTAAGAAAGTGTATTTATGCCCTACAGGAAGCCCTTACAGAGTCTACCTCCCTATC 905
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3193 AATTATGTAAGAAAGTGTATTTATGCCCTACAGGAAGCCCTTACAGAGTCTACCTCCCTATC 3252
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 906 CCAGCATCCCGACTCTTCCCCAAATTAAGGACCCCGCTTCAACCCAAATGGTCCAA 965
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3253 CCAGCATCCCGACTCTTCCCGCAATTAAGGACCCCGCTTCAACCCAAATGGTCCAA 3312
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 966 AAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCTAATTTCCCAATTAAGA 1025
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3313 AAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCTAATTTCCCAATTAAGA 3372
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1026 CCCTCCCAAGCAGTGGAGGAAGAAATTCGGCCAGCAGAGTGCATGTCGCTTTT 1085
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3373 CCCTCCCAAGCAGTGGAGGAAGAAATTCGGCCAGCAGAGTGCATGTCGCTTTT 3431
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1086 TCTCCAGACTTAAGCAATAAAGACAGACTTAGGTAAATCTCAGATAAATCTGATGG 1145
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3432 TCTCCAGACTTAAGCAATAAAGACAGACTTAGGTAAATCTCAGATAAATCTGATGG 3491
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1146 CTATATTTGTTTACAGGGTTAGACAATTTCTTTGATCTGACATGGAGATATATA 1205
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db 3492 CTATATTTGATGTTTTTACAAGGGTTAGACAATTTCTTTTGATCTGACATGGAGATATA-A 3550
QY 1206 TGTCACTGCTAAATCAGACACTAACCCCAATCAGAGAAAGTGCACCATAACTGCAGCCT 1265
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3551 TGTCACTGCTAAATCAGACACTAACCCCAATCAGAGAAAGTGCACCATAACTGCAGCCT 3610
QY 1266 GAGGTTTGGCGATCTCTGGTATCTCAGTCAGTCAATGATGATGATGATGATGATGATGATG 1325
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3611 GAGAGTTTGGCGATCTCTGGTATCTCAGTCAGTCAATG-ATAGGATGACACAG-AGG 3667
QY 1326 AAAGANAATGATTTCCCAAGCCAGCAGCAGTCCCACTCTCCAGTCTTASACCTCATTTGGGAC 1385
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3668 AAAGANAATGATTTCCCAAGCCAGCAGCAGTCCCACTCTCAGTCTTASACCTCATTTGG- 3724
QY 1386 ACAGAAATCAGTAACATGGGAGATTGGTCTGCAGACATTTGCTAACTTTGTGCTASAA 1445
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3725 ACACAGAAATCAGAAACAT-GGAGATTGGTCTGCAGACATTTGCTAACTTTGTGCTAGAA 3783
QY 1446 GGNCTAAGGAAACTAGGAAAGAAATCTAYGAATTTACTCAATGATCTCCACCATTAACACA 1505
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3784 GGACTAAGGAAACTAGGAAAG-AGTCTATGAATTTACTCAATGATGCTCCACCATTAACACA 3842
QY 1506 GGGGAAGGGAAGAAATCTTACTGCTCTTCTGGAGAGACTAAGGAGGCAATTCAGGAAGC 1565
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3843 -GGGAAGGGAAGAAATCTTACTGCTCTTCTGGAGAGACTAAGGAGGCAATTCAGGAAGC 3901
QY 1566 GTGCTCTCTGTGCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAACTTTATCA 1625
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3902 GTGCTCTCTGTGCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAACTTTATCA 3961
QY 1626 CTCAGTCAGTCGACACATTTAG-AAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAA 1684
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3962 CTCAGTCAGTCGACACATTTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAA 4021
QY 1685 ACTTAGAAACCTTATTGAACCTTGGCAACTCGSTTTTTTATATATAGAGATCAGAGGAGC 1744
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4022 ACTTAGAAACCTTATTGAACCTTGGCAACTCGSTTTTTTATATATAGAGATCAGAGGAGC 4081
QY 1745 AGCGGAAACAGGACAAACCGGATTTAAAAAAGGCCACCGCTTTAGTCAATGACCTCAGG 1804
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4082 AGCGGAAACAGGACAAACCGGATTTAAAAAAGGCCACCGCTTTAGTCAATGACCTCAGG 4141
QY 1805 CAAAGTGAATTTGGAGGCTCTGAAAAAGGAAAGCTGGGCAAAATGAATGCTTAATAGG 1864
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4142 CAAAGTGAATTTGGAGGCTCTGAAAAAGGAAAGCTGGGCAAAATGAATGCTTAATAGG 4201
QY 1865 GCTTGTCTCCAGTGGCTCTACAGGACACTTTTAAAAAGATTTGTCAGATAGAGTAAG 1924
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4202 GCTTGTCTCCAGTGGCTCTACAGGACACTTTTAAAAAAGATTTGTCAGATAGAGTAAG 4261
QY 1925 CGCGCCCTTCTGTCATGCCCTTTATTTCAAGGGAATCACTGGAGGCCCACTGCCCCAGG 1984
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4262 CGCGCCCTTCTGTCATGCCCTTTATTTCAAGGGAATCACTGGAGGCCCACTGCCCCAGG 4321
QY 1985 GGACAAAGGTCTTTTGAAGTCAGAGCCACTAAACAGATGATCCAGCAGCAGGACTGAGGG 2044
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4322 GGACAAAGGTCTCTCTGAGTCAGAGCCACTAAACAGATGATCCAGCAGCAGGACTGAGGG 4381
QY 2045 TGCTGGGCAAGCGGCATCCATGTCATCACTCCAGAGCCCTGGGTATGCTTTGACC 2104
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4382 TGCTGGGCAAGCGGCATCCATGTCATCACTCCAGAGCCCTGGGTATGCTTTGACC 4441
QY 2105 ATTGAGGGCCAGAAAGTTGCTCTCTGGACACTGGTGGCTCTTCTAGTCTTACTCTTC 2164
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4442 ATTGAGGGCCAGG-AGTTTGTCTCTTGGACACTGGTGGCTCTTCTAGTCTTACTCTTC 4500
QY 2165 TGTCCCGCAAACTGCTCCAGATCTGTCACTATTTCTGAGGGGGTCCNTAAGACGGGCA 2224
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4501 TGTCCCGCAAACTGCTCTCCAGATCTGTCACTA-TCTGAGGGGGTCC-TAAGACGGGCA 4558
QY 2225 GTCACTAGATCTTTTCCAGCCACTAAGTATGAACTGGGGAGCTTTATTTCTTTTCAC 2284
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4559 GTCACTAGATCTTTCTCCAGCCACTAAGTATG-ACTGGGGAGCTTTATTTCTTTTCAC 4616
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```













QY 3003 TTACTTTTRGCVGCCYRTTCAGAAACCTTGTCCATCAAGCCACCAAGCBCTCTTMAAT 3062  
Db 5263 TTACTTTTGGCGGCCCATTCAGAAACCTTGTGCCATCAAGCCACCAAGCGCTCTTCAAT 5322  
QY 3063 TTCTTCGCVACCTGTGGCTACAWGGTTTCCAAACSRARGCTCARCTCTGTCTCACAGCAG 3122  
Db 5323 TTCTTCGCTACTGTGGCTACATGGTTTCCAAACCAAGGCTCAACTCTGTCTCAGGCGAG 5382  
QY 3123 GTTAAATACTTAGBCTTAARATTATCCAAAGGCAACARGGCCCTCAGTGGAGAAVRYATC 3182  
Db 5383 GT- ---TACTTAGGGCTAAATATTCCAAAGGCCACAGGCACTCAGTGGAGAAACATC 5438  
QY 3183 CAGCTTATCTGGCTTATCTCATCYCAAAACCTTAAGCACTAAGRRTTCTCTGGC 3242  
Db 5439 CAGCCTATCTGGCTTATCTCATCCCCAAACCTTAAGCACTAAGGGGATTCCTTGGC 5498  
QY 3243 RTAAVAGGYTTTCGCCGAATATGGAATCCCGAGTGTGGCRAAATAGCAGGYCAATTAWA 3302  
Db 5499 GTAATAGGTTTCTGCCGAATATGGATT-CCGAGGTATGGCGNAATAGCCAGTCAATTAA 5557  
QY 3303 TACASTAATTAAGGAAATCTGAAAGCCAAATACCCATTTPARTPAAGATGGAYMCTGAAGY 3362  
Db 5558 TACACTAATTAAGGAAATCTGAAAGCCAAATACCCATTTAGTAAGATGGCACTGAAGT 5617  
QY 3363 MRAAGTGGCTTCCAGGCCCTTAAGAAGGCCCTTAAACCCAGYCCCACTGTAAGYTTG 3422  
Db 5618 AGAAGTGGCTTTCCAGGGCCC-----TAACAAGACCCCACTGTAAAGTTTG 5663  
QY 3423 CCAACRGGGCAAGACTTTTSTTYATAYRTACAGAAAAAACAAGAAVAGCTCTRGAGT 3482  
Db 5664 CCNAAGGGCAAGACTTTTCTCATATGTCAAG-AAAAACAGGNAATAGCTCTAGGAGT 5722  
QY 3483 CTTTACACAGRTCCRAGGAYGAGCTTGCAACYRTGGRYACCTGASTAAGGAAAYTGA 3542  
Db 5723 CCTTACACAGATCCGAGGGATGAGCTTGCAACTGTGGCATACCTGACTAAGGAAATTTA 5782  
QY 3543 TGTAGTGGCAAGGGTTGBCYTCATGTGTATGGGTAGTGGTGGCAGTAGCAGTYKTAGT 3602  
Db 5783 TGTAGTGGCAAGGGTTGACCTCATGTGTATGGGTAGTGGTGGCAGTAGCAGTCTTAGT 5842  
QY 3603 ATCTCAAGCAGTTAAATAATACAGGAGAGATCTTACTGTGTGGACATCTCATGAKGT 3662  
Db 5843 ATCTGAGCAGTTAAATAATACAGGAGAGATCTTACTGTGTGGACATCTCATGATGT 5902  
QY 3663 GAAYRGCACTACTGCTTAAAGGAGACTTGTGGCTGTGAGCAACACYGTFTTACTTAAATR 3722  
Db 5903 GAATGGCATACTCACTGTTAAAGGAGACTTGTGGCTGTGAGCAACACTGTTTACTTAAATG 5962  
QY 3723 TCAGGCTCTATTACTTGAAGGCCAGTGTGCACTGTGCACTTGTGCACTTTAAACCC 3782  
Db 5963 TCAGGCTCTATTACTTGAAGGCCAGTGTGCACTGTGCACTTGTGCACTTTAAACCC 6022  
QY 3783 AGYCNCAATTTCTTCAGACAATGAAGAAAGATAAATAACTGTCAACAARTAAATTC 3842  
Db 6023 AGCCACATTTCTTCAGACAATGAAGAAAGATAGAACATAACTGTCAACAARTAAATTC 6082  
QY 3843 TCAAAACCTATGCCACTCGAGGGGACCTTGTAGARGTTCCYTTGACTGTATCCYGACTTCA 3902  
Db 6083 TCAAAACCTATGCCACTCGAGGGGACCTTTTAGAGGTTCTCTTGAAGTCCCGACC-TCA 6141  
QY 3903 ACTTGTATCTGATGGAAGTTCTTGTAGAAAAAGGACTTCGAAAAAGYGGGGTATGAG 3962  
Db 6142 ACTTGTATCTGATGGAAGTTCTTGTAGAAAAAGGACTTTGAAAAAGTGGGGTATGAG 6201  
QY 3963 TGGTCAGTGAATAAGAAATATTTGAAAGTAATCCCTCACTCCAGGAACCTAGTGTCTYAGC 4022  
Db 6202 TGGTCAGCGAATAAGAAATATTTGAAAGTAATCCCTCACTCCAGGAACCTAGTGTCTYAGC 6261  
QY 4023 TRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTAGGAGAAAGAAAAAGGGYAAATA 4082  
Db 6262 TAGCAGAACTAATAGCCCTCACTCGGGCACTAGAATTAGGAGAAAGAAAAAGGGCAATA 6321

QY 4083 TATATACAGACTCTRAATATGCTVACCTAGTCTCCTCCTCCATGMRGCCAATATGARSAG 4142  
Db 6322 TATATACAGACTCTRAATATGCTVACCTAGTCTCCTCCTCCATGCCCATGCGAATATGGAAG 6381  
QY 4143 AAAGGGAATTCCTAACTTCYAGRGAAACACCTATCAMAATCAGAAAGCCATTAGGARAT 4202  
Db 6382 AAAGGGAATTCCTAACTTCYAGRGAAACACCTATCAAACATCAGGAAGCCATTAGGAAT 6441  
QY 4203 TATTATGGCWGTACAGAAACCTARAGAGTGGMGAGTCTTACACTGCGGGGTCAATCANA 4262  
Db 6442 TATTATGGCTGTATCAGAAACCTTAAGAGTGGGAGTCTTACACTGCGGGGTCAATCANA 6501  
QY 4263 AAGGAAAGAAAGGGAATASAGGAATTCGAAGCAKATATTCGAAGCAAAAGAGCTG 4322  
Db 6502 AAGGAAAGAAAGGGAATASAGGAATTCGAAGCAKATATTCGAAGCAAAAGAGCTG 6561  
QY 4323 CAAGCAGGACCTCCATTAGAAATGCTTATAAATCTTCCCTTACTAGTATAGGGTAAATCCCT 4382  
Db 6562 CAAGCAGGACCTCCATTAGAAATGCTTATAAATCAACCCCTAGTATAGGGTAAATCCCT 6621  
QY 4383 TCCGGGAAACCAAGCCCGAGTACTCAGCAGAGAAACAGAAATGGGGAACCTCAGGAG-C 4441  
Db 6622 TCCGGGAAACCAAGCCCGAGTACTCAGCAGAGAAACAGAAATGGGGAACCTCAGGAGAC 6681  
QY 4442 AGTTTTCCTCCCTCGGAGCGTGTAGCCACTGAAGAAAGGAAATACTTTTGGCTGCAACT 4501  
Db 6682 AGTTTTCCTCCCTCGGAGCGTGTAGCCACTGAAGAAAGGAAATACTTTTGGCTGCAACT 6741  
QY 4502 ATCCAAATGGAATTAATTTAAACCCCTTCATCAACCTTTTCACTTAGGCATCGATGACCC 4561  
Db 6742 AACCAATGGAATTAATTTAAACCCCTTCATCAACCTTTTCACTTAGGCATCGATGACCC 6801  
QY 4562 CATCARATGGCAATATCAATTTACTGGAACGAGCCCTTTTCAAACTATCAAGCARATA 4621  
Db 6802 CATCAGATGGCAATATCAATTTACTGGAACGAGCCCTTTTCAAACTATTTAAGCAGATA 6861  
QY 4622 KTCAGGGCTGTGAATGTGCAARAATAATATCCCTGCGCTTATGCGCAAGCTCCTTCA 4681  
Db 6862 GTCAGGGCTGTGAATGTGCAARAATAATATCCCTGCGCTTATGCGCAAGCTCCTTCA 6921  
QY 4682 GGARAACAAABAACAGGCGCATTTACCTGRARAARAATGCGCAACTGATTTTACCCAAAGC 4741  
Db 6922 GGARAACAAABAACAGGCGCATTTACCTGGAAGAAGCTGGAACCTGATTTTACCCAAAGC 6981  
QY 4742 CCAAACTCTCAGGATTTTCAAGTATCTACTAGTCTGGGTARATATCTTTCAGGGGTGGGCA 4801  
Db 6982 CCAAACTCTCAGGATTTTCAAGTATCTACTAGTCTGGGTARATATCTTTCAGGGGTGGGCA 7041  
QY 4802 AGGCTCTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACCTAGTTCATGAAA 4861  
Db 7042 AGGCTCTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACCTAGTTCATGAAA 7101  
QY 4862 TAAATCCAGATTCGAGCTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCAGG 4921  
Db 7102 TAAATCCAGATTCGAGCTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCAGG 7161  
QY 4922 CCAAGTAAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGGCCT 4981  
Db 7162 CCAAGTAAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGGCCT 7221  
QY 4982 GAAGGCCACAGTCTCAGGGAAGTTCGAGAAATCAATGAATGAATGAATGAATGAATGAAT 5041  
Db 7222 GAAGGCCACAGTCTCAGGGAAGTTCGAGAAATCAATGAATGAATGAATGAATGAATGAAT 7281  
QY 5042 AAAAGCAAAACCCAGGAAACCCACCTCACAATGSCCTGYTCTGTGTGCTATAGCTTAAAA 5101  
Db 7282 AAAAGCAAAACCCAGGAAACCCACCTCACAATGSCCTGYTCTGTGTGCTATAGCTTAAAA 7341  
QY 5102 GAATCTGCAACTTTCCCAAAAGCAGGACTTAGCCCATACCAATGCTGTATGAGGAGC 5161  
Db 7342 GAATCTGCAACTTTCCCAAAAGCAGGACTTAGCCCATACCAATGCTGTATGAGGAGC 7401  
QY 5162 CTTTCAATCAATGACCTTGTGCTTGACCCCAAGACGCAACTTAGTGTGACACATCAC 5221

Db 7402 CTTTCATAACCAATGACCTTGTGCTTGA-CCNAGACAGCCAACTTAGTTGAGACATCAC 7460  
QY 5222 CTCCTTAGCCAAATATCAACAAGTTCTTAAAAATTAACAAGAACCTTATCCCTGAGAGA 5281  
Db 7461 CTCCTTAGCCAAATATCAACAAGTTCTTAAAAATTAACAAGAACCTTATCCCTGAGAGA 7520  
QY 5282 GGGAAAAGAACTATTCACACCCWGTGACATGTTAGTCAAGTCCCTTCYCTCTAATTC 5341  
Db 7521 GGGAAAAGAACTATTCACACCTTGTGACATGGTATAGTCAAGTCCCTTCCTCTAATTC 7580  
QY 5342 CCCATCCCTAGATACATCTCTGGGAAGACCCCTACCCAGTCAATTTATATACCCCAACTGC 5401  
Db 7581 CCCATCCCTAGATACATCTCTGGGAAGACCCCTACCCAGTCAATTTATCTACCCCACTGC 7640  
QY 5402 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATACACTTGGAGTCAAAATCTCTGGATACT 5461  
Db 7641 GGTAAAGTGGCTGGAGTGGAGTCTTGGATACATACACTCGAGTCAAAACCTCTGGATACT 7700  
QY 5462 GCCAAAAGAACCTGAAAATCCAGGAGACACGCTAGCTATCTCTGTGAACCTCTAGAGGA 5521  
Db 7701 GCCAAAAGAACCTGAAAATCCAGGAGACACGCTAGCTATCTCTGTGAACCTCTAGAGGA 7760  
QY 5522 TTTGCGCTCTCTCTCAAAACAAACACGAGGAGAAAGTAACTAAATCATAAATCCCCCA 5581  
Db 7761 TTTGCGCTCTCTCTCAAAACAAACACGAGGAGAAAGTAACTAAATCATAAAT-ACCCA 7819  
QY 5582 TGGSCCTCCCTTATCATATTTTCTCTKTAATGTSTTTTACCCCTSTTTTCACTCTCACTG 5641  
Db 7820 TGGSCCTCCCTTATCATATTTTCTCTTACTGTCTTTTACCCCTCTTCTCACTCTCACTG 7879  
QY 5642 CACCCCTCCATCGCGCTGATGACCAAGTAGTCCCTTACCMAGAGTTTCTATGAGAA 5701  
Db 7880 CACCCCTCCATCGCGCTGATGACCAAGTAGTCCCTTACCAAGAGTTTCTATGAGAA 7939  
QY 5702 TGCAGCGTCCGGAATATTTGATGCCCATCTGATAGGAGTCTTTTAAAGGAAACCCCA 5761  
Db 7940 TGCAGCGTCCGGAATATTTGATGCCCATCTGATAGGAGTCTTTTAAAGGAAACCCCA 7999  
QY 5762 CTTTCACTGCCACACCCCATATGCCCGCAACTGCTATCACTGCGCACTCTTTGCAATGC 5821  
Db 8000 CTTTCACTGCCACACCCCATATGCCCGCAACTGCTATCACTGCGCACTCTTTGCAATGC 8059  
QY 5822 ATGCAATACTCATTTATGAGACAGGAAAATGATTAATCTAGTGTCTCTGAGGACTTG 5881  
Db 8060 ATGCAATACTCATTTATGAGACAGGAAAATGATTAATCTAGTGTCTCTGAGGACTTG 8119  
QY 5882 GAGTCACTGTCTGTTGCACTTACTTCAACCAACTGGTATGCTGATGGGGTGGAGTTC 5941  
Db 8120 GAGTCACTGTCTGTTGCACTTACTTCAACCATCTGGTATGCTGATGGGGTGGAGTTC 8179  
QY 5942 AAGATCAGGCAAGAGAAAACATGTAAGAAAGTAACTCTCCAACTCACCGGGTACATG 6001  
Db 8180 AAGATCAGGCAAGAGAAAACATGTAAGAAAGTAACTCTCCAACTCACCGGGTACATA 8239  
QY 6002 GCACCTTAGCCCTTCAAGGACTAGATCTCTCAAAAATACATGAACCCCTCGTACCC 6061  
Db 8240 GCAGCTCTAGCCCTTCAAGGACTAGATCTCTCAAAAATACATGAACCCCTCGTACCC 8299  
QY 6062 ATACTCGCTGGTAAAGCTTATTAATACCACTCTAGTGGCTCATGAGGTCTCGGCC 6121  
Db 8300 ATACTCGCTGGTAAAGCTTATTAATACCACTCTAGTGGCTCATGAGGTCTCGGCC 8359  
QY 6122 AAAACCTTACTTACTGTGGATATGCTCCCTCGAACTTTCARGCCATATGTTTCAATCC 6181  
Db 8360 AAAACCTTACTTACTGTGGATATGCTCCCTCGAACTTTCARGCCATATGTTTCAATCC 8419  
QY 6182 CTGTACTGAAATGGAACAACTTCAGCACAGAAATAAACACCACTTCGCTTTTATGATG 6241  
Db 8420 CTGTACTGAAATGGAACAACTTCAGCACAGAAATAAACACCACTTCGCTTTTATGATG 8479  
QY 6242 GACCTCTGTTTCCAAATSTGGAATAAACCCATACCTCAAAACCTTCACTGTGTAATAATTA 6301

Db 8480 GACCTCTGTTTCCAAATCTGGAATTAACCATACCTCAAAACCTCACCTGTGTGTAATAATTA 8539  
QY 6302 GCAATACTACATACAAACCAACTCCCAATGCAATGAGTGGGTAACTCTCCACACAAA 6361  
Db 8540 GCAATACTACATACAAACCAACTCCCAATGCAATGAGTGGGTAACTCTCCACACAAA 8599  
QY 6362 TAGTCTGCTTACCCCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCTGTTTGTGA 6421  
Db 8600 TAGTCTGCTTACCCCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCTGTTTGTGA 8659  
QY 6422 ATGCTCTTTCAGAAATCTATGTGCTCTCTCTCATCTTCTAGTGGCCCVATGRCATCTACA 6481  
Db 8660 ATGCTCTTTCAGAAATCTATGTGCTTCTCTCATCTTCTAGTGGCCCTATGACCATCTACA 8719  
QY 6482 CTGAAACAAGATTTATACATGTTATGTCTATCTAAGCCCGCAACAAAGAGTAGTACCCATTC 6541  
Db 8720 CTGAAACAAGATTTATACAATTTATGTCGTATCTAAGCCCGCAACAAAGAGTAGTACCCATTC 8779  
QY 6542 TTTCTTTTGTATAGGAGCAGGAGTCTAGGTGCACTAGGTACTGCGATTTGGCGGTATCA 6601  
Db 8780 TTTCTTTTGTATAGGAGCAGGAGTCTAGGTGCACTAGGTACTGCGATTTGGCGGTATCA 8839  
QY 6602 CAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTTAATGGGACATGGAACGG 6661  
Db 8840 CAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTTAATGGGACATGGAACGG 8899  
QY 6662 TGCAGGACTCCCTGGTCACTTTCGAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTTC 6721  
Db 8900 TGCAGGACTCCCTGGTCACTTTCGAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTTC 8959  
QY 6722 RAAATCGAAGAGCTTTAGACTGTCTAAACCGCTGAGAGAGGGGAACTGTTTATTTTAG 6781  
Db 8960 AAAATCGAAGAGCTTTAGACTGTCTAAACCGCTGAGAGAGGGGAACTGTTTATTTTAG 9019  
QY 6782 GGGAGAATGCTGTTATTTATGTTAATCAATCCGAATCTGCTACTGAGAAAGTTTAAAGAA 6841  
Db 9020 GGGAGAATGCTGTTATTTATGTTAATCAATCCGAATCTGCTACTGAGAAAGTTTAAAGAA 9079  
QY 6842 TTSAGATCGAATAACAACGAKAGAGAGCTTCGAAACACTGGAACCTCGGGGCTCC 6901  
Db 9080 TTSAGATCGAATAACAACGAKAGAGAGCTTCGAAACACTGGAACCTCGGGGCTCC 9139  
QY 6902 TCAGCCATAGGAGCTTCCCTTCTTGGACCTCTAGCAGCTTATTAATTTGC 6961  
Db 9140 TCAGCCATAGGAGCTTCCCTTCTTGGACCTCTTGGACCTCTTATTAATTTGC 9199  
QY 6962 TACTCTCTTGGACCTGATCTTTRACCTCTGTTTAACTTGTCTCTTCAGAAATCG 7021  
Db 9200 TACTCTCTTGGACCTGATCTTAACTCTCTTAACTTGTCTCTTCAGAAATCG 9259  
QY 7022 AAGCTGTAAAATCAAAATGGAGCCCAAGATGCAAGTCCAAAGACTAAGACTACCGCAGAC 7081  
Db 9260 AAGCTGTAAAATCAAAATGGAGCCCAAGATGCAAGTCCAAAGACTAAGACTACCGCAGAC 9319  
QY 7082 CCTTGGACCGCTGTATAGCCCAAGCTGATGTTAATGACATCAAAAGGACCCCTCTCTG 7141  
Db 9320 CCTTGGACCGCTGTATAGCCCAAGCTGATGTTAATGACATCAAAAGGACCCCTCTCTG 9379  
QY 7142 AGGAAATCTCAGCTGCAACCTCTACTAGCCGCCCAATTCAGAGGAAAGTGTAGAGCG 7201  
Db 9380 AGGAAATCTTAACTGCAACCTCTACTAGCCGCCCAATTCAGAGGAAAGTGTAGAGCG 9439  
QY 7202 GTSGTGGGCAACCTTCCCAACAGCACTTGTAGTTTCTGTTGAGATGGGGACTCAGAG 7261  
Db 9440 GTSGTGGGCAACCTTCCCAACAGCACTTGTAGTTTCTGTTGAGATGGGGACTCAGAG 9499  
QY 7262 ACAGGACTAGCTGGATTTCTTAGGCTGATTAAGAACTCCYTAAGCTTAGTGGGAAGTGA 7321  
Db 9500 ACAGGACTAGCTGGATTTCTTAGGCTGATTAAGAACTCCYTAAGCTTAGTGGGAAGTGA 9559  
QY 7322 CCACATCCACTTTTAAACACGGGCTTGCAACTTATGTCACACCTGACCAATCAGAGAGC 7381  
Db 9560 CCACATCCACTTTTAAACACGGGCTTGCAACTTATGTCACACCTGACCAATCAGAGAGC 9619

Qy	7382	TCACCTAAATGCTAATTAGGCAAGACAGGAGGTAAAGAAATAGCCAATCATYTTATTGCM	7441
Db	9620	TCACCTAAATGCTAATTAGGCAAGAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGCC	9679
Qy	7442	TGAGAGCACAGCAGGAGGACATGATCGGGATATAAACCCCAAGTTCGAGCCGGCAAC	7501
Db	9680	TGAGAGCACAGCAGGAGGACATGATCGGGATATAAACCCCAAGTTCGAGCTGGCAAC	9739
Qy	7502	GSCAACCCCTTTGGGTCCCTCTTTGTATGGAGCTCTGTTTCATGCTATTTCAC	7561
Db	9740	GSCAACCCCTTTGGGTCCCTCTTTGTATGGAGCTCTGTTTCATGCTATTTCAC	9799
Qy	7562	CTATTAAATCTTGCACTGCR	7582
Db	9800	CTATTAAATCTTGCACTGCA	9820

Search completed: August 2, 2006, 03:16:35  
Job time : 27475 secs



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2006, 15:08:51 ; Search time 5532 Seconds  
(without alignments)  
16841.079 Million cell updates/sec

Title: US-10-717-580-11

Perfect score: 7489.6

Sequence: 1 caacaatcggtataaacc.....tattaatacttgactgtgr 7582

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 9: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*
- 10: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*
- 11: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*
- 12: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10G\_PUBCOMB.seq:\*
- 13: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11A\_PUBCOMB.seq:\*
- 14: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11B\_PUBCOMB.seq:\*
- 15: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11C\_PUBCOMB.seq:\*
- 16: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11D\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	7489.6	100.0	7582	8	US-10-632-793-30	Sequence 30, Appl
2	7489.6	100.0	7582	8	US-10-717-580-11	Sequence 11, Appl
3	6448	86.1	56093	3	US-09-873-367C-81	Sequence 81, Appl
4	6448	86.1	56093	10	US-10-843-641A-81	Sequence 81, Appl
5	6436	85.9	10499	13	US-11-028-539-3	Sequence 3, Appl
6	6124.6	81.8	8523	3	US-09-854-867-21	Sequence 21, Appl
7	6124.6	81.8	8523	10	US-10-786-970A-21	Sequence 21, Appl
8	3738.4	49.9	6394	10	US-10-450-763-20014	Sequence 20014, A
9	3431	45.8	4349	10	US-10-450-763-12279	Sequence 12279, A
10	3309.6	44.2	7466	10	US-10-450-763-4430	Sequence 4430, Ap
11	3105.4	41.5	177175	15	US-11-121-086-79	Sequence 79, Appl
12	3105.4	41.5	189993	15	US-11-121-086-78	Sequence 78, Appl
13	3049.8	40.7	22436	3	US-09-997-722-148	Sequence 148, App
14	2948	39.4	139573	12	US-10-506-513-3	Sequence 3, Appl
15	2948	39.4	156416	10	US-10-461-862-12	Sequence 12, Appl
16	2932	39.1	1514	10	US-10-450-763-3413	Sequence 3413, Ap
17	2889.6	38.6	2930	3	US-09-902-535-1	Sequence 1, Appli

# ALIGNMENTS

## RESULT 1

US-10-632-793-30  
; Sequence 30, Application US/10632793  
; Publication No. US20040048298A1  
; GENERAL INFORMATION:  
; APPLICANT: PARAMHOS-BACCALA, Glaucia  
; APPLICANT: MALLET, Francois  
; APPLICANT: VOISET, Cecile  
; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN  
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT  
; FILE REFERENCE: 110048  
; CURRENT APPLICATION NUMBER: US/10/632,793  
; PRIOR FILING DATE: 2003-08-04  
; PRIOR APPLICATION NUMBER: US/09/869,927  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: PCT/FR00/00144  
; PRIOR FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: FR 99/00888  
; PRIOR FILING DATE: 1999-01-21  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 30  
; LENGTH: 7582  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (198)..(198)  
; OTHER INFORMATION: n = a or g or c or t/u  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (307)..(307)  
; OTHER INFORMATION: n = a or g or c or t/u  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (355)..(355)  
; OTHER INFORMATION: n = a or g or c or t/u  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1309)..(1309)

18 2888.6 38.6 2946 6 US-10-114-893-134 Sequence 134, App  
19 2888.6 38.6 2946 6 US-10-016-249-3 Sequence 3, Appli  
20 2734.4 36.5 3831 10 US-10-450-763-7531 Sequence 7531, Ap  
21 2688.6 35.9 2938 8 US-10-632-793-21 Sequence 21, Appl  
22 2688.6 35.9 2938 9 US-10-717-580-2 Sequence 2, Appli  
23 2613.2 34.9 3372 8 US-10-632-793-28 Sequence 28, Appl  
24 2613.2 34.9 3372 9 US-10-717-580-9 Sequence 9, Appli  
25 2550.4 34.1 2599 13 US-11-028-539-1 Sequence 1, Appli  
26 2541.6 33.9 4535 10 US-10-450-763-12009 Sequence 12009, A  
27 2368.4 31.6 8279 10 US-10-450-763-12278 Sequence 12278, A  
28 2282.4 30.5 2784 13 US-11-028-539-4 Sequence 4, Appli  
29 2251.8 30.1 2372 8 US-10-632-793-29 Sequence 29, Appl  
30 2251.8 30.1 2372 9 US-10-717-580-10 Sequence 10, Appl  
31 2241 29.9 8294 10 US-10-450-763-20013 Sequence 20013, A  
32 2220.4 29.6 161334 6 US-10-087-192-730 Sequence 730, App  
33 2216 29.6 46340 13 US-11-028-539-66 Sequence 66, Appl  
34 2204.6 29.4 2782 8 US-10-632-793-26 Sequence 26, Appl  
35 2204.6 29.4 2782 9 US-10-717-580-7 Sequence 7, Appli  
36 2201.4 29.4 2782 8 US-10-133-036-1 Sequence 1, Appli  
37 2201.4 29.4 3633 10 US-10-821-234-543 Sequence 543, App  
38 2168 28.9 8399 3 US-09-854-867-26 Sequence 26, Appl  
39 2168 28.9 8399 10 US-10-786-970A-26 Sequence 26, Appl  
40 2131.4 28.5 2575 9 US-10-717-580-14 Sequence 14, Appl  
41 2049.4 27.4 3205 10 US-10-450-763-8040 Sequence 8040, Ap  
42 2032 27.1 2055 13 US-11-028-539-22 Sequence 22, Appl  
43 2026.6 27.1 2585 9 US-10-717-580-13 Sequence 13, Appl  
44 1982.2 26.5 3867 10 US-10-450-763-4425 Sequence 4425, Ap  
45 1982.2 26.5 3867 10 US-10-450-763-12268 Sequence 12268, A

```
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1331)..(1331)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2213)..(2213)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2398)..(2398)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3787)..(3787)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4115)..(4115)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4261)..(4261)
; OTHER INFORMATION: n = a or g or c or t/u
US-10-632-793-30

Query Match 100.0%; Score 7489.6; DB 8; Length 7582;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAACAAATCGGGATATAAACCAGGCAATTCGAGCTGGCAACAGAGCCCCCTTTGGGTCC 60
DB 1 CAACAAATCGGGATATAAACCAGGCAATTCGAGCTGGCAACAGAGCCCCCTTTGGGTCC 60

QY 61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTTCACTTAAATCTTCAACTGCA 120
DB 61 CTTCCCTTTGTATGGGAGCTGTTTTCATGCTATTTTCACTTAAATCTTCAACTGCA 120

QY 121 CTCCTTCGTGCTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTACCGCTCCACCACTGC 180
DB 121 CTCCTTCGTGCTCATGTTTCTTACGGCTCGAGCTGAGCTTTTGTCTACCGCTCCACCACTGC 180

QY 181 TGTGTTGCCACCAACCGCANACTGCGCTGACTCCCATCCCTCTGGATCTCGAGGGTGC 240
DB 181 TGTGTTGCCACCAACCGCANACTGCGCTGACTCCCATCCCTCTGGATCTCGAGGGTGC 240

QY 241 CGCTGTGCTCCTGATCCAGGARGCGCCCATTCGCGCTCCCAATTTGGGCTTAAAGGCTGC 300
DB 241 CGCTGTGCTCCTGATCCAGGARGCGCCCATTCGCGCTCCCAATTTGGGCTTAAAGGCTGC 300

QY 301 CATTTGNCCTGCAAGGCTTAAGTGCCTGGGTTTGTCTTAATTTGAGCTGAACACTANTCACT 360
DB 301 CATTTGNCCTGCAAGGCTTAAGTGCCTGGGTTTGTCTTAATTTGAGCTGAACACTANTCACT 360

QY 361 GGGTTCCATGGTTCTCTTCTGTGACCCAGCGCTTCTTAATAAACAATAAACAATAAACA 420
DB 361 GGGTTCCATGGTTCTCTTCTGTGACCCAGCGCTTCTTAATAAACAATAAACAATAAACA 420

QY 421 TGGCCCCAAGATTCATTCCTTGGAAATCCGTGAGGSCAAGCAATCCAGGCTCAGAGATAC 480
DB 421 TGGCCCCAAGATTCATTCCTTGGAAATCCGTGAGGSCAAGCAATCCAGGCTCAGAGATAC 480

QY 481 GARGCTTGGCACCACATCTTGGAAAGCGGCTGCTACCTCTTGGAAAGTGGTTTCAACCAATC 540
DB 481 GARGCTTGGCACCACATCTTGGAAAGCGGCTGCTACCTCTTGGAAAGTGGTTTCAACCAATC 540

QY 541 TTGGGAGCTCTGTGAGCAAGGACCCCGGTGACATTTTGGGACCAAGGACGATCC 600
DB 541 TTGGGAGCTCTGTGAGCAAGGACCCCGGTGACATTTTGGGACCAAGGACGATCC 600

QY 601 MAAGTGCATGGAAACGTTCCCGCAAGACAAAACCCCTTAAGAGCTATTTCTGGARAAT 660
DB 601 MAAGTGCATGGAAACGTTCCCGCAAGACAAAACCCCTTAAGAGCTATTTCTGGARAAT 660

601 MAAGTGCATGGAAACGTTCCCGCAAGACAAAACCCCTTAAGAGCTATTTCTGGARAAT 660
661 TGGGAMCAATTTGACCCCTCAGACACTAAGAAAGAAACGACTTATATTTCTTGGCAGTGCC 720
661 TGGGAMCAATTTGACCCCTCAGACACTAAGAAAGAAACGACTTATATTTCTTGGCAGTGCC 720
721 GCCTGGCACTCCTGAGGGAAGTATATAATTAACACCATCTTACAGCTAGACATCTTTTG 780
721 GCCTGGCACTCCTGAGGGAAGTATATAATTAACACCATCTTACAGCTAGACATCTTTTG 780
781 TAGAAAAGGCAAAATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 840
781 TAGAAAAGGCAAAATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 840
841 CTCACAAATTTATGTAATAAAGTGTATTTATGCCCCTACAGGAAGCCCTTCAGAGTCTACCTCC 900
841 CTCACAAATTTATGTAATAAAGTGTATTTATGCCCCTACAGGAAGCCCTTCAGAGTCTACCTCC 900
901 CTATCCAGCATCCCCGACTCCTTCCCAATTAAGAGACCCCTTCAACCCCAATGG 960
901 CTATCCAGCATCCCCGACTCCTTCCCAATTAAGAGACCCCTTCAACCCCAATGG 960
961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCTAATATTTCCCAAT 1020
961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCTAATATTTCCCAAT 1020
1021 TATGACCCCTCCCAAGCAGTGGGAGGAGAAATTCGGCCCAAGCAGAGTGATGTCYT 1080
1021 TATGACCCCTCCCAAGCAGTGGGAGGAGAAATTCGGCCCAAGCAGAGTGATGTCYT 1080
1081 TTTTTCCTCCAGACTTAAAGCAATTAAGCAAGTCTTAAATTTCTCAGATAATTCCT 1140
1081 TTTTTCCTCCAGACTTAAAGCAATTAAGCAAGTCTTAAATTTCTCAGATAATTCCT 1140
1141 GATGCTATATTTGTTTACAGGGTTAGGACAAATTTCTTGTATCTGACATGAGAGAT 1200
1141 GATGCTATATTTGTTTACAGGGTTAGGACAAATTTCTTGTATCTGACATGAGAGAT 1200
1201 ATATATGTCACCTGCTTAAATCAGACACTAACCCCAATGAGAGAGTGCCACCAATCTGC 1260
1201 ATATATGTCACCTGCTTAAATCAGACACTAACCCCAATGAGAGAGTGCCACCAATCTGC 1260
1261 AGCTGAGGTTTGGGATCTCTGATCTCAGTCAGTCAATGATGATGATGATGATGATGAT 1320
1261 AGCTGAGGTTTGGGATCTCTGATCTCAGTCAGTCAATGATGATGATGATGATGATGAT 1320
1321 GAAGGAAAGANAAATGATTTCCCAAGGCGAGCAGTTCAGTCTTASACCTCATTTG 1380
1321 GAAGGAAAGANAAATGATTTCCCAAGGCGAGCAGTTCAGTCTTASACCTCATTTG 1380
1381 GGGACACAGAAATCAGTAAACATGGGAGATTTGGTCTGCAGACATTTGCTTAACTTGTGTC 1440
1381 GGGACACAGAAATCAGTAAACATGGGAGATTTGGTCTGCAGACATTTGCTTAACTTGTGTC 1440
1441 TASAAGGCTAAGGAAACTAGGAAATCTAAGAAATCTAAGAAATCTAAGAAATCTAAGAAAT 1500
1441 TASAAGGCTAAGGAAACTAGGAAATCTAAGAAATCTAAGAAATCTAAGAAATCTAAGAAAT 1500
1501 ACACAGGCGGAGGGAAGAAATCTTCTGAGAGAGTTCAGGAGAGTTCAGGAGGAGGAGT 1560
1501 ACACAGGCGGAGGGAAGAAATCTTCTGAGAGAGTTCAGGAGAGTTCAGGAGGAGGAGT 1560
1561 GAAGCGTGCCTCTCTGTACCTGACTCTTCTGAAAGGCAACTAATCTTAAAGCGTAAAGTT 1620
1561 GAAGCGTGCCTCTCTGTACCTGACTCTTCTGAAAGGCAACTAATCTTAAAGCGTAAAGTT 1620
1621 TATCACTCAGTCAGTCAGACATTTAGAAAACCTTCAAAAGCTTCGCCCTAGGCCCGGAG 1680
1621 TATCACTCAGTCAGTCAGACATTTAGAAAACCTTCAAAAGCTTCGCCCTAGGCCCGGAG 1680
1681 CAAAACCTTAGAAACCTTATTTGCAACCTTGGCAACCTTGGTCTTAAATAGAGATCAGGAG 1740
1681 CAAAACCTTAGAAACCTTATTTGCAACCTTGGCAACCTTGGTCTTAAATAGAGATCAGGAG 1740
```

Qy	1741	GAGCAGCGGAAACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCCCT	1800
Db	1741	GAGCAGCGGAAACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCCCT	1800
Qy	1801	CAGGCAAGTGACATCTTGAGGCTCTGGAAAGGGAAAAAGCTGGGCAAAATGGAATGCCTAA	1860
Db	1801	CAGGCAAGTGACATCTTGAGGCTCTGGAAAGGGAAAAAGCTGGGCAAAATGGAATGCCTAA	1860
Qy	1861	TAGGGCTGCTTCCAGTGCGETCTACAAGGACATTTTAAAAAAGATTTGTCTCAGTAGAAG	1920
Db	1861	TAGGGCTGCTTCCAGTGCGETCTACAAGGACATTTTAAAAAAGATTTGTCTCAGTAGAAG	1920
Qy	1921	TAAAGCGCCCTTCGTCCATGCCCTTATTTCAAGGGAATCACATGGAAGGCCCACTGCCCC	1980
Db	1921	TAAAGCGCCCTTCGTCCATGCCCTTATTTCAAGGGAATCACATGGAAGGCCCACTGCCCC	1980
Qy	1981	CAGGGGACAAAGGCTTTTGTAGTCAGAAAGCCACTTAACAGATGATCCAGCAGCAGGACTG	2040
Db	1981	CAGGGGACAAAGGCTTTTGTAGTCAGAAAGCCACTTAACAGATGATCCAGCAGCAGGACTG	2040
Qy	2041	AGGCTGCTGGGGCAAGCGCATCCCATGCCATCACCCTCACAGAGCCCTTGGGTATGCTT	2100
Db	2041	AGGCTGCTGGGGCAAGCGCATCCCATGCCATCACCCTCACAGAGCCCTTGGGTATGCTT	2100
Qy	2101	GACCATTTGAGGGCCAGGAAGTTGCTCTCTGGACATCTGCTGCGGTCTTCTTAGTCTTACT	2160
Db	2101	GACCATTTGAGGGCCAGGAAGTTGCTCTCTGGACATCTGCTGCGGTCTTCTTAGTCTTACT	2160
Qy	2161	CTTCTGTCCGGACAACTGTCCTCCAGATCTGTCACTATTTCTGAGGGGGTCCNTAAGACG	2220
Db	2161	CTTCTGTCCGGACAACTGTCCTCCAGATCTGTCACTATTTCTGAGGGGGTCCNTAAGACG	2220
Qy	2221	GGCAGTCACTAGATACCTTTTCCAGCCACTAAGTTATGAACCTGGGGAGCTTTATCTTTT	2280
Db	2221	GGCAGTCACTAGATACCTTTTCCAGCCACTAAGTTATGAACCTGGGGAGCTTTATCTTTT	2280
Qy	2281	TCACATGCTTTTCTAAATTATGCTTGAAGGCCCACTACCTTGTGTAGGGAGAGACATTTCTA	2340
Db	2281	TCACATGCTTTTCTAAATTATGCTTGAAGGCCCACTACCTTGTGTAGGGAGAGACATTTCTA	2340
Qy	2341	GCAAAAGAGGGGCCATTATACACCTGAAATAGGGAAGGAACACCCGTTTGTGTGTTGCC	2400
Db	2341	GCAAAAGAGGGGCCATTATACACCTGAAATAGGGAAGGAACACCCGTTTGTGTGTTGCC	2400
Qy	2401	CCTGCTTTGGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAGGAACAATATGACGAGC	2460
Db	2401	CCTGCTTTGGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAGGAACAATATGACGAGC	2460
Qy	2461	CAAGAATGCCGCTCTGTTTCAAGTTAAAATGAAGGATCCACTTCTCTTCCCTACCAAA	2520
Db	2461	CAAGAATGCCGCTCTGTTTCAAGTTAAAATGAAGGATCCACTTCTCTTCCCTACCAAA	2520
Qy	2521	GGCAGTACCCCTTCAGACCCAAAGGCCCAACAAAGGATTTCAAAGAATTTGTTAAGGACTTAA	2580
Db	2521	GGCAGTACCCCTTCAGACCCAAAGGCCCAACAAAGGATTTCAAAGAATTTGTTAAGGACTTAA	2580
Qy	2581	AAGCCCAAGGCTTAGTAAACATGATCACTACTCCCTGCAGTAAATCCGTAGTGGATTGAG	2640
Db	2581	AAGCCCAAGGCTTAGTAAACATGATCACTACTCCCTGCAGTAAATCCGTAGTGGATTGAG	2640
Qy	2641	GAGGSCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAATCAATGG	2700
Db	2641	GAGGSCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTAATCAATGG	2700
Qy	2701	AGGCCGTTGCTTTTATAACCAAGCTGTACTAGCCCTTACTAGTGMVTTCCAAATAC	2760
Db	2701	AGGCCGTTGCTTTTATAACCAAGCTGTACTAGCCCTTACTAGTGMVTTCCAAATAC	2760
Qy	2761	CAGAGGAACAGAGTGGTTTACASTCCTGACCTTAAAGGATGCTTCTTCTGATCCCTTG	2820
Db	2761	CAGAGGAACAGAGTGGTTTACASTCCTGACCTTAAAGGATGCTTCTTCTGATCCCTTG	2820

Qy	2821	TACATCCTGACTCTCAATTCCTGTGTTTGCCTTTGAGATACCTTCAAAACCCARCATCTCAAC	2880
Db	2821	TACATCCTGACTCTCAATTCCTGTGTTTGCCTTTGAGATACCTTCAAAACCCARCATCTCAAC	2880
Qy	2881	TCACCTGGACTRTTTTAAACCCCAAGGGTTTCAGGGATAGYCCCCATCTATTTTGGCCAGGCGAT	2940
Db	2881	TCACCTGGACTRTTTTAAACCCCAAGGGTTTCAGGGATAGYCCCCATCTATTTTGGCCAGGCGAT	2940
Qy	2941	TAGCCCAAGACTTGGAGYCAATYMTCATACCTGGACACTCTTGTGCTTTCRGTAAGTGGGATG	3000
Db	2941	TAGCCCAAGACTTGGAGYCAATYMTCATACCTGGACACTCTTGTGCTTTCRGTAAGTGGGATG	3000
Qy	3001	ATTTTACTTTTTRGCGVCCYRTTTCAGAAACCTTGTGCCATCAAGCCACCAAGCRCTCTTMA	3060
Db	3001	ATTTTACTTTTTRGCGVCCYRTTTCAGAAACCTTGTGCCATCAAGCCACCAAGCRCTCTTMA	3060
Qy	3061	ATTTTCCTCGVACCTGTGGCTACAWGGTTTCCAAASARARGTTCARCTCTGTGCTCACAGC	3120
Db	3061	ATTTTCCTCGVACCTGTGGCTACAWGGTTTCCAAASARARGTTCARCTCTGTGCTCACAGC	3120
Qy	3121	AGGTTAAATACTTAGGRCCTAARATTAATCCAAAGGCCACARGGCCCTCAGTGAGGAAYRYA	3180
Db	3121	AGGTTAAATACTTAGGRCCTAARATTAATCCAAAGGCCACARGGCCCTCAGTGAGGAAYRYA	3180
Qy	3181	TCCAGCCCTATACCTGCTTATCCTCATCYCAAAACCCCTTAAGCAACTAAGHGRRTTCTCTTG	3240
Db	3181	TCCAGCCCTATACCTGCTTATCCTCATCYCAAAACCCCTTAAGCAACTAAGHGRRTTCTCTTG	3240
Qy	3241	GCRTAAYAGGYTTCGCCGAATATGGANTCCCCAGGTGTGCCRAAATAGCCAGGYCATTA	3300
Db	3241	GCRTAAYAGGYTTCGCCGAATATGGANTCCCCAGGTGTGCCRAAATAGCCAGGYCATTA	3300
Qy	3301	WATACASTAATAAGGAAACTCAGAAAGCCCAATACCCATTTARTTAAGATGGAYAMCTGAA	3360
Db	3301	WATACASTAATAAGGAAACTCAGAAAGCCCAATACCCATTTARTTAAGATGGAYAMCTGAA	3360
Qy	3361	GYMRAAGTGGCTTTCAGGCCCTTAAAGAGGCTTTAAACCCAAAGYCCCAAGTGTTAAGYT	3420
Db	3361	GYMRAAGTGGCTTTCAGGCCCTTAAAGAGGCTTTAAACCCAAAGYCCCAAGTGTTAAGYT	3420
Qy	3421	TGCCAACRGGGCAAGACTTTTSTTYATVHTACAGAAAAACAGAAAYAGCTCTRGGGA	3480
Db	3421	TGCCAACRGGGCAAGACTTTTSTTYATVHTACAGAAAAACAGAAAYAGCTCTRGGGA	3480
Qy	3481	GTCCTTACACAGRTCCRAGGGAAGCTTGCAACCCYRTGGCRVACCTGASTAAGGAAAYT	3540
Db	3481	GTCCTTACACAGRTCCRAGGGAAGCTTGCAACCCYRTGGCRVACCTGASTAAGGAAAYT	3540
Qy	3541	GATGTAGTGCAAAAGGGTTGRCYTCATTTGTTTAYGGGTAGTGTGGCAGTAGCAGTYKTA	3600
Db	3541	GATGTAGTGCAAAAGGGTTGRCYTCATTTGTTTAYGGGTAGTGTGGCAGTAGCAGTYKTA	3600
Qy	3601	GTATCTGAAGCAGTTAAATAATACAGGGPAGAGATCTTACTGTGTGGACATCTCATGAK	3660
Db	3601	GTATCTGAAGCAGTTAAATAATACAGGGPAGAGATCTTACTGTGTGGACATCTCATGAK	3660
Qy	3661	GTGAAYRGCATACTCACTGCTTAAAGGAGACTTGTGCTGTGCAGACAACYGTTTACTTAAA	3720
Db	3661	GTGAAYRGCATACTCACTGCTTAAAGGAGACTTGTGCTGTGCAGACAACYGTTTACTTAAA	3720
Qy	3721	TRTCAGGCTCTATTAATCTTGAARGGCCAGTGTGCRACCTGTGCACTTGTGCAACTCTTAAAC	3780
Db	3721	TRTCAGGCTCTATTAATCTTGAARGGCCAGTGTGCRACCTGTGCACTTGTGCAACTCTTAAAC	3780
Qy	3781	CCAGYCNCAATTTCTCCAGACANTGAAGAAAGATARAAYATACTGTCAACAAPTAAAT	3840
Db	3781	CCAGYCNCAATTTCTCCAGACANTGAAGAAAGATARAAYATACTGTCAACAAPTAAAT	3840
Qy	3841	TCTCAAAACCTATGCCACTCGAGGGACCTTTTAGARGTTCCYTTGACTGATCCYGACCTT	3900
Db	3841	TCTCAAAACCTATGCCACTCGAGGGACCTTTTAGARGTTCCYTTGACTGATCCYGACCTT	3900
Qy	3901	CAACTTGTATCTAGTGGGAAGTTTCCTTTGTGTAAGAAAGGACCTTCGAAAAAGYGGGTATGCG	3960

Db 3901 CAACCTTGATCTAGTGAAGTTCCTTTGTAGAAAAGGACTTCGAAAAGYGGGTATGC 3960  
Qy 3961 AGTGTGCTAGTGAATATGGAATATTTGAAAGTAAATCCCTCACTCAGGAACTAGTGCTTAA 4020  
Db 3961 AGTGTGCTAGTGAATATGGAATATTTGAAAGTAAATCCCTCACTCAGGAACTAGTGCTTAA 4020  
Qy 4021 GCTRCGAGAACTAATAGCCYTCAYTKGGGCACTAGAAATTTAGGAGAAAGGAAAGGAA 4080  
Db 4021 GCTRCGAGAACTAATAGCCYTCAYTKGGGCACTAGAAATTTAGGAGAAAGGAAAGGAA 4080  
Qy 4081 TATATATACAGACTCTRTARTATGCTTAACTAGTCTNTCCATGMRGCAATATGSAR 4140  
Db 4081 TATATATACAGACTCTRTARTATGCTTAACTAGTCTNTCCATGMRGCAATATGSAR 4140  
Qy 4141 AGAAGGGAATTCCTAACTTCYVAGRGAAACCTATCAMACATCAGGAAGCCATTAGGAR 4200  
Db 4141 AGAAGGGAATTCCTAACTTCYVAGRGAAACCTATCAMACATCAGGAAGCCATTAGGAR 4200  
Qy 4201 ATTATTATYGGCWGTACAGAAACCTPARAGAGTGGWAGTCTTACACTGCYGGGTATCA 4260  
Db 4201 ATTATTATYGGCWGTACAGAAACCTPARAGAGTGGWAGTCTTACACTGCYGGGTATCA 4260  
Qy 4261 NAAAGGAAGRAAGGGAATASAGRGAAATGCAAGCAKATATTTGAAGCMAAAGAGC 4320  
Db 4261 NAAAGGAAGRAAGGGAATASAGRGAAATGCAAGCAKATATTTGAAGCMAAAGAGC 4320  
Qy 4321 TGCAAGGAGGACCCCTCCATTTAGAAATGCTTTAACTTCCCTTAGTATAGGGTAATCC 4380  
Db 4321 TGCAAGGAGGACCCCTCCATTTAGAAATGCTTTAACTTCCCTTAGTATAGGGTAATCC 4380  
Qy 4381 CTTCCGGGAAACCAAGCCCACTACTCAGCAGAGAAACAGATGGGNAACCTCAGGAG 4440  
Db 4381 CTTCCGGGAAACCAAGCCCACTACTCAGCAGAGAAACAGATGGGNAACCTCAGGAG 4440  
Qy 4441 CAGTTTTCTCCCTCGGACGGTGTAGCCACTGAAGGGAATACTTTTGCCTGCAAC 4500  
Db 4441 CAGTTTTCTCCCTCGGACGGTGTAGCCACTGAAGGGAATACTTTTGCCTGCAAC 4500  
Qy 4501 TATCCAATGGAATTTACTTAAACCCCTTCATCAAACTTTCACTAGGCATCGATAGCAC 4560  
Db 4501 TATCCAATGGAATTTACTTAAACCCCTTCATCAAACTTTCACTAGGCATCGATAGCAC 4560  
Qy 4561 CCATCARATGGCAATCAATTTACTTGGACAGGCCCTTTTCAAACTATCAAGCARAT 4620  
Db 4561 CCATCARATGGCAATCAATTTACTTGGACAGGCCCTTTTCAAACTATCAAGCARAT 4620  
Qy 4621 AKTCAAGGCTGTGAATGTGCCARAAATTAATCCCTGCTCTYATCGCAAGCTCTTC 4680  
Db 4621 AKTCAAGGCTGTGAATGTGCCARAAATTAATCCCTGCTCTYATCGCAAGCTCTTC 4680  
Qy 4681 AGGARAACAAARAACAGGCCATTACCTGRARAARACTGGCAACTGATTTTACCCACAAG 4740  
Db 4681 AGGARAACAAARAACAGGCCATTACCTGRARAARACTGGCAACTGATTTTACCCACAAG 4740  
Qy 4741 CCCAAACCTCAGGATTTTCAAGTATCTACTAGTCTGGTARATATCTTCAOGGTTGGCA 4800  
Db 4741 CCCAAACCTCAGGATTTTCAAGTATCTACTAGTCTGGTARATATCTTCAOGGTTGGCA 4800  
Qy 4801 RAGGCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTATGAA 4860  
Db 4801 RAGGCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTATGAA 4860  
Qy 4861 ATAATTTCCAGATTCGACTTCCCGAGGCTTACAGAGTGAACAATAGCCCTCTTTCCAG 4920  
Db 4861 ATAATTTCCAGATTCGACTTCCCGAGGCTTACAGAGTGAACAATAGCCCTCTTTCCAG 4920  
Qy 4921 GCCAGATTAACCGAGGATTTCCAGGCTTATAGGTATACGATATCACTTACACTGGCC 4980  
Db 4921 GCCAGATTAACCGAGGATTTCCAGGCTTATAGGTATACGATATCACTTACACTGGCC 4980  
Qy 4981 TGAAGGCCACAGTCTCTCAGGAAGGTTCAGAAAAATGAATGAAAYACTCAAAGGACATCTA 5040

Db 4981 TGAAGGCCACAGTCTCTCAGGGAAGGTTCGAGAAAAATGAATGAAAVACTCAAAGGACATCTA 5040  
Qy 5041 AAAAGGCAAAACCCAGGAAACCCACCTCACATGGCTGTCTGTGCTCTATAGCCTTAAAA 5100  
Db 5041 AAAAGGCAAAACCCAGGAAACCCACCTCACATGGCTGTCTGTGCTCTATAGCCTTAAAA 5100  
Qy 5101 AGAATCTGCAACTTTTCCCAAAAAACAGGACTTAGCCCATACGAAATGCTGTATGGAAG 5160  
Db 5101 AGAATCTGCAACTTTTCCCAAAAAACAGGACTTAGCCCATACGAAATGCTGTATGGAAG 5160  
Qy 5161 CCTTCATAACCAATGACTTGTGCTTGAACCAAGACAGCCAACTTAGTTCAGACATCA 5220  
Db 5161 CCTTCATAACCAATGACTTGTGCTTGAACCAAGACAGCCAACTTAGTTCAGACATCA 5220  
Qy 5221 CCTCTTAGCCAAATATCAAGTTCCTTAAAAACATTTAAGGAACCTATCCCTGAGAG 5280  
Db 5221 CCTCTTAGCCAAATATCAAGTTCCTTAAAAACATTTAAGGAACCTATCCCTGAGAG 5280  
Qy 5281 AGGAAAAAGAACTATTTCCACCCWMTGACATGGTATTAGTCAAGTCCCTTCCTTAAT 5340  
Db 5281 AGGAAAAAGAACTATTTCCACCCWMTGACATGGTATTAGTCAAGTCCCTTCCTTAAT 5340  
Qy 5341 CCCCATCCCTAGATACATCTCTGGGAAGACCCCTACCCAGTCAATTTATYTACCCCACTG 5400  
Db 5341 CCCCATCCCTAGATACATCTCTGGGAAGACCCCTACCCAGTCAATTTATYTACCCCACTG 5400  
Qy 5401 CGGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAACTCTGGATAC 5460  
Db 5401 CGGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGAGTCAAACTCTGGATAC 5460  
Qy 5461 TGCCAAAGGAACCTGAAATCCAGGAGAACCGCTAGCTATTTCTGTGAACTCTTAGAGG 5520  
Db 5461 TGCCAAAGGAACCTGAAATCCAGGAGAACCGCTAGCTATTTCTGTGAACTCTTAGAGG 5520  
Qy 5521 ATTTGGGCTGCTCTTCAAAACCAACAGGAGGAAGTAATAAAATCATAAATCCCC 5580  
Db 5521 ATTTGGGCTGCTCTTCAAAACCAACAGGAGGAAGTAATAAAATCATAAATCCCC 5580  
Qy 5581 ATGSGCTCCCTTATCATATTTTCTCTKTASTGTTTCTTACCTCTCTCACT 5640  
Db 5581 ATGSGCTCCCTTATCATATTTTCTCTKTASTGTTTCTTACCTCTCTCACT 5640  
Qy 5641 GCACCCCTCATGCGCTGTATGACCACTAGTCTCCCTYACCMAGAGTTTCTATGGAGA 5700  
Db 5641 GCACCCCTCATGCGCTGTATGACCACTAGTCTCCCTYACCMAGAGTTTCTATGGAGA 5700  
Qy 5701 ATGCAAGCTCCGGAATATTTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCC 5760  
Db 5701 ATGCAAGCTCCGGAATATTTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCC 5760  
Qy 5761 ACCTTCACTGCCACACCCCATATGCCCCGCACTGCTATCACTGCGCACTCTTTGCATG 5820  
Db 5761 ACCTTCACTGCCACACCCCATATGCCCCGCACTGCTATCACTGCGCACTCTTTGCATG 5820  
Qy 5821 CATGCAATATCTTATTTGGACAGGAAAAATGATTAATCTTAGTTCCTGAGGACTT 5880  
Db 5821 CATGCAATATCTTATTTGGACAGGAAAAATGATTAATCTTAGTTCCTGAGGACTT 5880  
Qy 5881 GGAGTCACTGTCTGTGGACTTACTTCAACCAAACTGGTATGCTGTATGGGGGTGGAGTT 5940  
Db 5881 GGAGTCACTGTCTGTGGACTTACTTCAACCAAACTGGTATGCTGTATGGGGGTGGAGTT 5940  
Qy 5941 CAGATCAGGCAAGGAAAAACATGTAAAAAGAGTAATCTCCCACTCACCSGGGTACAT 6000  
Db 5941 CAGATCAGGCAAGGAAAAACATGTAAAAAGAGTAATCTCCCACTCACCSGGGTACAT 6000  
Qy 6001 GGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCTCTCGTACC 6060  
Db 6001 GGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCTCTCGTACC 6060  
Qy 6061 CATACTCGCTGGTAAGCTATTTTAAATACCACTCTCTGAGGCTCCATGAGGTCTCGGCC 6120  
Db 6061 CATACTCGCTGGTAAGCTATTTTAAATACCACTCTCTGAGGCTCCATGAGGTCTCGGCC 6120

QY 6121 CAAAACTTAACTGTTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATC 6180  
DB |||||||  
QY 6121 CAAAACTTAACTGTTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATC 6180  
DB |||||||  
QY 6181 CTTGTACCTGAAACAATGGAACAACCTTCAGCACAGAAATAAACAACCACTTCGGTTTTAGTA 6240  
DB |||||||  
QY 6181 CTTGTACCTGAAACAATGGAACAACCTTCAGCACAGAAATAAACAACCACTTCGGTTTTAGTA 6240  
DB |||||||  
QY 6241 GGACCTCTTGTGTTCCAAATSTGGAATAAACCATACTCAAACTCACTGTGTGTAATAATTT 6300  
DB |||||||  
QY 6241 GGACCTCTTGTGTTCCAAATSTGGAATAAACCATACTCAAACTCACTGTGTGTAATAATTT 6300  
DB |||||||  
QY 6301 AGCAATACATACATACCAACCACTCCCAATGCATCAGGTGGTAACTCTCCCAACAA 6360  
DB |||||||  
QY 6301 AGCAATACATACATACCAACCACTCCCAATGCATCAGGTGGTAACTCTCCCAACAA 6360  
DB |||||||  
QY 6361 ATAGTCTGCTACCTCAGGAATATTTTTGCTGTGTGACCTCAGCCCTATGTTCTTTG 6420  
DB |||||||  
QY 6361 ATAGTCTGCTACCTCAGGAATATTTTTGCTGTGTGACCTCAGCCCTATGTTCTTTG 6420  
DB |||||||  
QY 6421 AATGGCTCTTCAAGATCTATGTGCTTCTCTCAATCTTAGTGCCCTCCATGRCATCTAC 6480  
DB |||||||  
QY 6421 AATGGCTCTTCAAGATCTATGTGCTTCTCTCAATCTTAGTGCCCTCCATGRCATCTAC 6480  
DB |||||||  
QY 6481 ACTGAACAAGATTTATACAGTTATGTTCATATCTAAGCCCGGCAAAAGAGATGCCATTT 6540  
DB |||||||  
QY 6481 ACTGAACAAGATTTATACAGTTATGTTCATATCTAAGCCCGGCAAAAGAGATGCCATTT 6540  
DB |||||||  
QY 6541 CTTCTCTTTGTTATAGGACGAGGTGCTAGGTGCACTAGGTACTGGCAATTCGGCGTATC 6600  
DB |||||||  
QY 6541 CTTCTCTTTGTTATAGGACGAGGTGCTAGGTGCACTAGGTACTGGCAATTCGGCGTATC 6600  
DB |||||||  
QY 6601 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG 6660  
DB |||||||  
QY 6601 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG 6660  
DB |||||||  
QY 6661 GTCCGCACTCCCTGTGTCACCTTGCAAGATCACTTAATCTCCCTAGCAGAGTAGTCCTT 6720  
DB |||||||  
QY 6661 GTCCGCACTCCCTGTGTCACCTTGCAAGATCACTTAATCTCCCTAGCAGAGTAGTCCTT 6720  
DB |||||||  
QY 6721 CRAAATCGAAGAGCTTTAGACTGCTAACCGCTGAGAGGGGGACCTGTTATTTTA 6780  
DB |||||||  
QY 6721 CRAAATCGAAGAGCTTTAGACTGCTAACCGCTGAGAGGGGGACCTGTTATTTTA 6780  
DB |||||||  
QY 6781 GGGGAAGAATGCTGTTATTTATGTTAATCAATCGGAATCGTCACTGAGAAAGTTTAAAGAA 6840  
DB |||||||  
QY 6781 GGGGAAGAATGCTGTTATTTATGTTAATCAATCGGAATCGTCACTGAGAAAGTTTAAAGAA 6840  
DB |||||||  
QY 6841 ATTCAGATCGAATACAACTAGCAGAGAGCTTCGAAACACATGGACCCCTGGGCGCTC 6900  
DB |||||||  
QY 6841 ATTCAGATCGAATACAACTAGCAGAGAGCTTCGAAACACATGGACCCCTGGGCGCTC 6900  
DB |||||||  
QY 6901 CTCAGCCATGATGCCCTGGATTTCCCTCTTAGGACCTCTAGCAGCTATAATATG 6960  
DB |||||||  
QY 6901 CTCAGCCATGATGCCCTGGATTTCCCTCTTAGGACCTCTAGCAGCTATAATATG 6960  
DB |||||||  
QY 6961 CTACTCTCTTTGGACCCCTGATCTTTTACCTCTGTTAACTTTGCTCTTTCCAGATC 7020  
DB |||||||  
QY 6961 CTACTCTCTTTGGACCCCTGATCTTTTACCTCTGTTAACTTTGCTCTTTCCAGATC 7020  
DB |||||||  
QY 7021 GAAGCTGTAAACTACAAATGGAGCCCAAGATGCAAGTCAAGATCTACCGCAGA 7080  
DB |||||||  
QY 7021 GAAGCTGTAAACTACAAATGGAGCCCAAGATGCAAGTCAAGATCTACCGCAGA 7080  
DB |||||||  
QY 7081 CCCCTGGACCGGCTGTTAGCCCAAGATCTGATGTTAATGATCAAAAGGCAACCCCTCT 7140  
DB |||||||  
QY 7081 CCCCTGGACCGGCTGTTAGCCCAAGATCTGATGTTAATGATCAAAAGGCAACCCCTCT 7140  
DB |||||||  
QY 7141 GAGGAATCTAGCTGCAAACTCTACTACGCCCCCAATTCAGCAGGAGAGTTAGAGC 7200  
DB |||||||  
QY 7141 GAGGAATCTAGCTGCAAACTCTACTACGCCCCCAATTCAGCAGGAGAGTTAGAGC 7200  
DB |||||||

QY 7201 GGTSGTCGGCAACCTCCCAACAGCACTTAGGTTTTCTGTTGAGATGGGGACTGAGA 7260  
DB |||||||  
QY 7201 GGTSGTCGGCAACCTCCCAACAGCACTTAGGTTTTCTGTTGAGATGGGGACTGAGA 7260  
DB |||||||  
QY 7261 GACAGACTAGCTGATTCCTAGCTGATTAAGAAATCCYTAAGCCTAGSTGGGAAGGTG 7320  
DB |||||||  
QY 7261 GACAGACTAGCTGATTCCTAGCTGATTAAGAAATCCYTAAGCCTAGSTGGGAAGGTG 7320  
DB |||||||  
QY 7321 ACCACATCCACTTTTAAACACAGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAG 7380  
DB |||||||  
QY 7321 ACCACATCCACTTTTAAACACAGGGGCTTGCAACTTAGYTCACACCTGACCAATCAGAGAG 7380  
DB |||||||  
QY 7381 CTCACATAAATGCTAATTAGGCAAAAGACAGGAGGTAAAGAAATAGCAATCATTTATTC 7440  
DB |||||||  
QY 7381 CTCACATAAATGCTAATTAGGCAAAAGACAGGAGGTAAAGAAATAGCAATCATTTATTC 7440  
DB |||||||  
QY 7441 MTGAGAGCACAGCAGGAGGCAATGATCGGGATATAAACCCAAAGTTCGAGCCGGCAA 7500  
DB |||||||  
QY 7441 MTGAGAGCACAGCAGGAGGCAATGATCGGGATATAAACCCAAAGTTCGAGCCGGCAA 7500  
DB |||||||  
QY 7501 CGGCAACCCCTTTGGGTCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCAC 7560  
DB |||||||  
QY 7501 CGGCAACCCCTTTGGGTCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCAC 7560  
DB |||||||  
QY 7561 TCTATTAAATCTTGCACTGCR 7582  
DB |||||||  
QY 7561 TCTATTAAATCTTGCACTGCR 7582  
DB |||||||

## RESULT 2

## US-10-717-580-11

; Sequence 11, Application US/10717580

; Publication No. US20040176314A1

; GENERAL INFORMATION:

; APPLICANT: BESEME, Frederic

; APPLICANT: BLOND, Jean-Luc

; APPLICANT: BOUTON, Olivier

; APPLICANT: MANDRAND, Bernard

; APPLICANT: MALLET, Francois

; APPLICANT: PERRON, Herve

; TITLE OF INVENTION: ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE DIS-

; TITLE OF INVENTION: WITH PREGNANCY DISORDERS

; FILE REFERENCE: 105045

; CURRENT APPLICATION NUMBER: US/10/717,580

; CURRENT FILING DATE: 2003-11-21

; PRIOR APPLICATION NUMBER: US/09/446,024A

; PRIOR FILING DATE: 1999-12-16

; PRIOR APPLICATION NUMBER: PCT/FR98/01442

; PRIOR FILING DATE: 1998-07-06

; PRIOR APPLICATION NUMBER: FR 97/08815

; PRIOR FILING DATE: 1997-07-07

; NUMBER OF SEQ ID NOS: 53

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 11

; LENGTH: 7582

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: (198)..(198)

; OTHER INFORMATION: n = any nucleotide

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: (307)..(307)

; OTHER INFORMATION: n = any nucleotide

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: (355)..(355)

; OTHER INFORMATION: n = any nucleotide

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: (1309)..(1309)

; OTHER INFORMATION: n = any nucleotide

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: (1309)..(1309)

; OTHER INFORMATION: n = any nucleotide

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1331)..(1331)
; OTHER INFORMATION: n = any nucleotide
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2213)..(2213)
; OTHER INFORMATION: n = any nucleotide
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2398)..(2398)
; OTHER INFORMATION: n = any nucleotide
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3787)..(3787)
; OTHER INFORMATION: n = any nucleotide
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4115)..(4115)
; OTHER INFORMATION: n = any nucleotide
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4261)..(4261)
; OTHER INFORMATION: n = any nucleotide
;
US-10-717-580-11

Query Match      100.0%; Score 7489.6; DB 9; Length 7582;
Beat Local Similarity 100.0%; Pred. No. 0;
Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAACAATCGGATATAAACCAGGCGATTCGAGCTGGCAACAGCAGCCGCCCTTTGGGTCC 60
DB 1 CAACAATCGGATATAAACCAGGCGATTCGAGCTGGCAACAGCAGCCGCCCTTTGGGTCC 60

QY 61 CTTCCCTTTGATGGGAGCTGTTTCATGCTATTTCACTCTATTAATAATCTTGCAACTGCA 120
DB 61 CTTCCCTTTGATGGGAGCTGTTTCATGCTATTTCACTCTATTAATAATCTTGCAACTGCA 120

QY 121 CTCCTTCGTGTCATGTTCTTAACGGCTCGAGCTGAGCTGAGCTTGGCTCCTGAGGCTGC 180
DB 121 CTCCTTCGTGTCATGTTCTTAACGGCTCGAGCTGAGCTTGGCTCCTGAGGCTGC 180

QY 181 TGTGTTGCCACACCCGACACCTGCGCTGACTCCCATCCCTCTGATCTCTGAGGCTGC 240
DB 181 TGTGTTGCCACACCCGACACCTGCGCTGACTCCCATCCCTCTGATCTCTGAGGCTGC 240

QY 241 CGCTGTGCTCTGATCCAGCGAGCGCCCAATGGCGCTCCCAATTTGGGCTAAAGGCTTGC 300
DB 241 CGCTGTGCTCTGATCCAGCGAGCGCCCAATGGCGCTCCCAATTTGGGCTAAAGGCTTGC 300

QY 301 CATTTGTCCTGACCGGCTAAGTGCCTGGGTTTGTCTTAATGAGCTGAACACTANTCACT 360
DB 301 CATTTGTCCTGACCGGCTAAGTGCCTGGGTTTGTCTTAATGAGCTGAACACTANTCACT 360

QY 361 GGGTTCATGTTCTCTGTGACCCAGCGCTTCTATAKAACATATAACACTTACCACA 420
DB 361 GGGTTCATGTTCTCTGTGACCCAGCGCTTCTATAKAACATATAACACTTACCACA 420

QY 421 TGGCCCAAGATTCCATTCTCTGGAATCCGTGAGGSCAACGAACTCCAGCTCAGAGATAC 480
DB 421 TGGCCCAAGATTCCATTCTCTGGAATCCGTGAGGSCAACGAACTCCAGCTCAGAGATAC 480

QY 481 GARGCTTGCCACATCTTGGAAAGCGGCTGTGTACGRTCTTGGAAAGTGGTTTCAACCATC 540
DB 481 GARGCTTGCCACATCTTGGAAAGCGGCTGTGTACGRTCTTGGAAAGTGGTTTCAACCATC 540

QY 541 TTGGGAGCTCTGTGACAGAGACCCCGGTACATTTTGGGACCCMSRACGGACATCC 600
DB 541 TTGGGAGCTCTGTGACAGAGACCCCGGTACATTTTGGGACCCMSRACGGACATCC 600

QY 601 MAAGTGATGGGAAACGTTCCCGCGAAGACAAAAACGCCCCCTTAAGAGCTATTTCTGGARAT 660
DB 601 MAAGTGATGGGAAACGTTCCCGCGAAGACAAAAACGCCCCCTTAAGAGCTATTTCTGGARAT 660
```

```
QY 661 TGGAMCAATTTTGACCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCC 720
DB 661 TGGAMCAATTTTGACCTCAGACACTAAGAAAGAAACGACTTATATTCTTCTGCAGTGCC 720

QY 721 GCCTGGCAGCTCCTGAGGGAAGTATAAATTAACACCATCTTACAGCTAGACYTCTTTTG 780
DB 721 GCCTGGCAGCTCCTGAGGGAAGTATAAATTAACACCATCTTACAGCTAGACYTCTTTTG 780

QY 781 TAGAAAAGGCAAAATGGAGTGAAGTGCCATAGTAACAACTTTCTTTTCATTAAAGACAA 840
DB 781 TAGAAAAGGCAAAATGGAGTGAAGTGCCATAGTAACAACTTTCTTTTCATTAAAGACAA 840

QY 841 CTCACAAATTATGTAAGTCTGATTATATGCCCTACAGGAAGCCTTACAGAGCTTACCTCC 900
DB 841 CTCACAAATTATGTAAGTCTGATTATATGCCCTACAGGAAGCCTTACAGAGCTTACCTCC 900

QY 901 CTATCCAGCATCCCGACTCTCTTCCCACMYTAATAAGGACCCCTTCAACCCAAATGG 960
DB 901 CTATCCAGCATCCCGACTCTCTTCCCACMYTAATAAGGACCCCTTCAACCCAAATGG 960

QY 961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAAACCAAGAGTGCCAAATATTCGCCAAT 1020
DB 961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAAACCAAGAGTGCCAAATATTCGCCAAT 1020

QY 1021 TATGACCCCTCCCAAGCAGTGGGAGGAGAAATTCGSCCCAGCCAGAGTGCATGTGCVT 1080
DB 1021 TATGACCCCTCCCAAGCAGTGGGAGGAGAAATTCGSCCCAGCCAGAGTGCATGTGCVT 1080

QY 1081 TTTTYYTCTCCAGACTTTAAAGCAAAATAAAAACAGACTTTAGGTAAATTTCTCAGATAAYCCT 1140
DB 1081 TTTTYYTCTCCAGACTTTAAAGCAAAATAAAAACAGACTTTAGGTAAATTTCTCAGATAAYCCT 1140

QY 1141 GATGGCTATATGRTGTTTTTCAAGGGTTAGGACAAATCTTTTGATCTGACATGGAGAGAT 1200
DB 1141 GATGGCTATATGRTGTTTTTCAAGGGTTAGGACAAATCTTTTGATCTGACATGGAGAGAT 1200

QY 1201 ATATATGTCAGTCTAAATCAGACACTAACCCCAATGAGAGAAAGTGCCACCATAACTGC 1260
DB 1201 ATATATGTCAGTCTAAATCAGACACTAACCCCAATGAGAGAAAGTGCCACCATAACTGC 1260

QY 1261 AGCTGAGRGTTTTGGCGATCTCTGATATCTCAGTCAAGTCAATGGATANGATGACAACA 1320
DB 1261 AGCTGAGRGTTTTGGCGATCTCTGATATCTCAGTCAAGTCAATGGATANGATGACAACA 1320

QY 1321 GAAGAAAGANAATGATTTCCCAAGCCAGCAGAGCTTCCAGCTTASACCTCATTTG 1380
DB 1321 GAAGAAAGANAATGATTTCCCAAGCCAGCAGAGCTTCCAGCTTASACCTCATTTG 1380

QY 1381 GGGACACAGAATCAGTAACATGGGAGATTGCTGTCGAGACATTTGCTAACTTGTGTGC 1440
DB 1381 GGGACACAGAATCAGTAACATGGGAGATTGCTGCTGAGACATTTGCTAACTTGTGTGC 1440

QY 1441 TASAAGGACTAAGGAAAACTASGAAGAARTCTAYGAATTAATCAATGATGTCCACCATTA 1500
DB 1441 TASAAGGACTAAGGAAAACTASGAAGAARTCTAYGAATTAATCAATGATGTCCACCATTA 1500

QY 1501 ACACAGGGAAGGGAAGAAAAATCTCTGCTTTCTGGAGAGACTTAAGGAGGCAATTGAG 1560
DB 1501 ACACAGGGAAGGGAAGAAAAATCTCTGCTTTCTGGAGAGACTTAAGGAGGCAATTGAG 1560

QY 1561 GAAAGGTGCTCTCTGTCACTGACTCTTCTGAGAGGCCAACTTAATCTTAAAGCGTAAGTT 1620
DB 1561 GAAAGGTGCTCTCTGTCACTGACTCTTCTGAGAGGCCAACTTAATCTTAAAGCGTAAGTT 1620

QY 1621 TATCACTCAGTCACTGAGACATTTAGAAAAAACTTCAAAAAGTCTGCGGTAGGCCCGGAG 1680
DB 1621 TATCACTCAGTCACTGAGACATTTAGAAAAAACTTCAAAAAGTCTGCGGTAGGCCCGGAG 1680

QY 1681 CAAAACCTTAGAAAAACCTTATTTGGAACCTTGGCAACCTTGGTGTGTTTTTATAATAGATCAGGAG 1740
DB 1681 CAAAACCTTAGAAAAACCTTATTTGGAACCTTGGCAACCTTGGTGTGTTTTTATAATAGATCAGGAG 1740
```



1741 GAGCAGCGGAAACAGGACAAACGGGATTAATAAAGGAGCCACCGCTTTAGTCATGACCT 1800  
1741 GAGCAGCGGAAACAGGACAAACGGGATTAATAAAGGAGCCACCGCTTTAGTCATGACCT 1800  
1801 CAGGCAAGTGGACTTTGGAGGCTCTGGAAAGGAAAGCTGGGCAAAATGAAATGGCTAA 1860  
1801 CAGGCAAGTGGACTTTGGAGGCTCTGGAAAGGAAAGCTGGGCAAAATGAAATGGCTAA 1860  
1861 TAGGGCTTGTCTCCAGTGGGCTTACAAGGACACTTTAAAAAGATTTGTCAGTAGAAG 1920  
1861 TAGGGCTTGTCTCCAGTGGGCTTACAAGGACACTTTAAAAAGATTTGTCAGTAGAAG 1920  
1921 TAAGCGCCCTCTCGTCATGCCCTTATTTCAAGGGAATCACTGGAAAGCCCACTGCC 1980  
1921 TAAGCGCCCTCTCGTCATGCCCTTATTTCAAGGGAATCACTGGAAAGCCCACTGCC 1980  
1981 CAGGGCAAGAGTCTTTTGAAGCAGAACCACTAACAGATGATCCAGCAGCAGACTG 2040  
1981 CAGGGCAAGAGTCTTTTGAAGCAGAACCACTAACAGATGATCCAGCAGCAGACTG 2040  
2041 AGGGTGCCTGGGCAAGCGCCATCCCATGCCATCACTCCCTCACAGAGCCCTGGGTATGCTT 2100  
2041 AGGGTGCCTGGGCAAGCGCCATCCCATGCCATCACTCCCTCACAGAGCCCTGGGTATGCTT 2100  
2101 GACCATTTGAGGCGCAGGAGTGTCTCTGGGACACTGTGGCGGTCTTCTTAGTCTTACT 2160  
2101 GACCATTTGAGGCGCAGGAGTGTCTCTGGGACACTGTGGCGGTCTTCTTAGTCTTACT 2160  
2161 CTTCTGTGCCGGAACAAGTCTCTCCAGATCTGTCACTATTTCTGAGGGGGTCCNTAAGAGC 2220  
2161 CTTCTGTGCCGGAACAAGTCTCTCCAGATCTGTCACTATTTCTGAGGGGGTCCNTAAGAGC 2220  
2221 GGCAAGTCACTAGATATCTTTTCCAGCACTAAGTATGAACTGGGAGCTTTATTTCTTT 2280  
2221 GGCAAGTCACTAGATATCTTTTCCAGCACTAAGTATGAACTGGGAGCTTTATTTCTTT 2280  
2281 TCACATGCTTTTCTAATATGCTTGAAGCCCACTACCTTGTAGGAGAGACATTTCTA 2340  
2281 TCACATGCTTTTCTAATATGCTTGAAGCCCACTACCTTGTAGGAGAGACATTTCTA 2340  
2341 GCAAAAGCAGGGGCCATTAACCTGAACATAGGAGAGGAACACCGTTTCTGTGCTGCC 2400  
2341 GCAAAAGCAGGGGCCATTAACCTGAACATAGGAGAGGAACACCGTTTGTGTGCTGCC 2400  
2401 CTTGCTTGAAGAGGAATTAATCCTGAAGTCTGGGCAACAGAGGAATATGAGCAGC 2460  
2401 CTTGCTTGAAGAGGAATTAATCCTGAAGTCTGGGCAACAGAGGAATATGAGCAGC 2460  
2461 CAAAGAAATCCCGTCTCTCAAGTTAACTAAGGATTTCCACTTCTTCCCTACCAAA 2520  
2461 CAAAGAAATCCCGTCTCTCAAGTTAACTAAGGATTTCCACTTCTTCCCTACCAAA 2520  
2521 GGCAGTACCCCTCAGACCAAGGCCCAACAGGATTTCCAAAGATTTGTTAAGGACTTAA 2580  
2521 GGCAGTACCCCTCAGACCAAGGCCCAACAGGATTTCCAAAGATTTGTTAAGGACTTAA 2580  
2581 AAGCCCAAGGCTTAGTAAACCAATGATTAATCCCTGCAAGTATTCCTGATGATGAG 2640  
2581 AAGCCCAAGGCTTAGTAAACCAATGATTAATCCCTGCAAGTATTCCTGATGATGAG 2640  
2641 GAGGCAAGAACCCAGTGGAGGAGTGTAGTGAAGATCTCAGGATTAATCAATGG 2700  
2641 GAGGCAAGAACCCAGTGGAGGAGTGTAGTGAAGATCTCAGGATTAATCAATGG 2700  
2701 AGGCGCTGTCTTTTATACCCAGCTGTACCTAGCCCTTATCTGTGATGATTTCCCAATAC 2760  
2701 AGGCGCTGTCTTTTATACCCAGCTGTACCTAGCCCTTATCTGTGATGATTTCCCAATAC 2760  
2761 CAGAGGAAGCAGAGTGGTATTAACSTCTGACCTTGAAGATGCTTTCTTCTGATCCCTG 2820  
2761 CAGAGGAAGCAGAGTGGTATTAACSTCTGACCTTGAAGATGCTTTCTTCTGATCCCTG 2820  
2821 TACATCCTGACTCTCAATTTCTTGTGCTTTGCTTTGAAGATACTTCAAAACCCATCTCAAC 2880

2821 TACATCCTGACTCTCAATTTCTTGTGCTTTGAAGATACTTCAAAACCCATCTCAAC 2880  
2881 TCACCTGGAGCTTTTACCCCAAGGTTTCAGGATAGYCCCATCTATTTTGGCCAGGCAAT 2940  
2881 TCACCTGGAGCTTTTACCCCAAGGTTTCAGGATAGYCCCATCTATTTTGGCCAGGCAAT 2940  
2941 TAGCCCAAGACTTTGAGYCAATYMTCACTGCTGACACTCTTGTCTTCTTCTGATGATG 3000  
2941 TAGCCCAAGACTTTGAGYCAATYMTCACTGCTGACACTCTTGTCTTCTTCTGATGATG 3000  
3001 ATTTTACTTTTTRGCGCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAGGCTCTTMA 3060  
3001 ATTTTACTTTTTRGCGCYRTTCAGAAACCTTGTGCCATCAAGCCACCCAGGCTCTTMA 3060  
3061 ATTTTCTCGCYACCTCTGCTACAWGGTTTCCAAACSAARAGCTCACTCTGCTCACAGC 3120  
3061 ATTTTCTCGCYACCTCTGCTACAWGGTTTCCAAACSAARAGCTCACTCTGCTCACAGC 3120  
3121 AGGTTAAATACCTTAGGCTTAARATTTTCCAAAGCCACCAAGGCTCAGTGAGGAAAT 3180  
3121 AGGTTAAATACCTTAGGCTTAARATTTTCCAAAGCCACCAAGGCTCAGTGAGGAAAT 3180  
3181 TCCAGCCTATCTGCTTATCTCATCYCAAAACCTTAAAGCAACTAAGGRTTCTTCTG 3240  
3181 TCCAGCCTATCTGCTTATCTCATCYCAAAACCTTAAAGCAACTAAGGRTTCTTCTG 3240  
3241 GRTTAAAGGTTTCTGCCGAATGATTTCCAGGTTGGGCAAAATAGCCAGYCATTA 3300  
3241 GRTTAAAGGTTTCTGCCGAATGATTTCCAGGTTGGGCAAAATAGCCAGYCATTA 3300  
3301 WATACASTAATTAAGGAACTCAGAAAGCAATACCCATTTARTAAAGTGGAYAMCTGAA 3360  
3301 WATACASTAATTAAGGAACTCAGAAAGCAATACCCATTTARTAAAGTGGAYAMCTGAA 3360  
3361 GYMAAGTGGCTTTCAGGCTTAAAGAGGCTTAAACCCAAAGYCCCAAGTCTTAAAGYT 3420  
3361 GYMAAGTGGCTTTCAGGCTTAAAGAGGCTTAAACCCAAAGYCCCAAGTCTTAAAGYT 3420  
3421 TGCACACRGGGCAAGACTTTTSTTTATVATTCACAGAAACCAACAGAAATAGCTTGG 3480  
3421 TGCACACRGGGCAAGACTTTTSTTTATVATTCACAGAAACCAACAGAAATAGCTTGG 3480  
3481 GTCTTTACACAGTCTCAGGAGGAGCTTTCAGAACCTTTCAGTAAAGGAAAT 3540  
3481 GTCTTTACACAGTCTCAGGAGGAGCTTTCAGAACCTTTCAGTAAAGGAAAT 3540  
3541 GATGATGCGCAAGGTTGTCATTTGTTTAYGGTATGTTGTCAGTGGAGTCTTAAAGYT 3600  
3541 GATGATGCGCAAGGTTGTCATTTGTTTAYGGTATGTTGTCAGTGGAGTCTTAAAGYT 3600  
3601 GTATCTGAGCAGTAAATATAACAGGAGAGATCTTCTGCTGAGGACATCTCATGAK 3660  
3601 GTATCTGAGCAGTAAATATAACAGGAGAGATCTTCTGCTGAGGACATCTCATGAK 3660  
3661 GTGAAATGCACTCACTGCTTAAAGGAGCTTGTGCTGCTCAGAACCTTGTGCTTAA 3720  
3661 GTGAAATGCACTCACTGCTTAAAGGAGCTTGTGCTGCTCAGAACCTTGTGCTTAA 3720  
3721 TRTCAAGCTTATTAATTGAARGGCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTTAA 3780  
3721 TRTCAAGCTTATTAATTGAARGGCCAGTCTGCTGCTGCTGCTGCTGCTGCTTAA 3780  
3781 CCAGT/CNCACTTCTTCCAGCAATGAGAAAGATARAAYATACTGTCAACAAATTAAT 3840  
3781 CCAGT/CNCACTTCTTCCAGCAATGAGAAAGATARAAYATACTGTCAACAAATTAAT 3840  
3841 TCTCAAACTATGCCACTCGAGGAGCTTGTAGAGTTCCTTGTGCTGATCCYGACCTT 3900  
3841 TCTCAAACTATGCCACTCGAGGAGCTTGTAGAGTTCCTTGTGCTGATCCYGACCTT 3900  
3901 CAACTTGTATCTGATGGAAGTTTCTTGTAGAAAGGACTTTCGAAAGYGGGATGCTC 3960

Db 3901 CAACCTTGATATGATGGAAGTTCCCTTTGTAGAAAAAGGACTTCGAAAAGYGGGTATGC 3960  
Qy 3961 AGTGTCTAGTATGAATATGGAATATTTGAAAGTAATCCCTCTACTCAGGAACTAGTGCTTAA 4020  
Db 3961 AGTGTCTAGTATGAATATTTGAAAGTAATCCCTCTACTCAGGAACTAGTGCTTAA 4020  
Qy 4021 GCTRCAGAACTAATAGCCYTCAYTKGGGCACCTAGAAATTTAGGAGAAAGAAAAGGYAAA 4080  
Db 4021 GCTRCAGAACTAATAGCCYTCAYTKGGGCACCTAGAAATTTAGGAGAAAGAAAAGGYAAA 4080  
Qy 4081 TATATATACAGACTCTRTARTATGCTTACCTAGTCTNTCCATGCCCATGMRGCAATATGSAR 4140  
Db 4081 TATATATACAGACTCTRTARTATGCTTACCTAGTCTNTCCATGCCCATGMRGCAATATGSAR 4140  
Qy 4141 AGAAAGGGAATTCCTAACTTCYVAGRGAAACACCTATCAMACATCAGGAAGCCATTAGGAR 4200  
Db 4141 AGAAAGGGAATTCCTAACTTCYVAGRGAAACACCTATCAMACATCAGGAAGCCATTAGGAR 4200  
Qy 4201 ATTATATATGCGWGTACAGAAACCTTARAGAGGTGGWAGTCTTACACTGCGGGGTCACTCA 4260  
Db 4201 ATTATATATGCGWGTACAGAAACCTTARAGAGGTGGWAGTCTTACACTGCGGGGTCACTCA 4260  
Qy 4261 NAAAGGAAAGAAAGGGAATASAAAGRGAAATGCGCAAGCAKATATTTGAAGCMAAAAGAGC 4320  
Db 4261 NAAAGGAAAGAAAGGGAATASAAAGRGAAATGCGCAAGCAKATATTTGAAGCMAAAAGAGC 4320  
Qy 4321 TGCAAGGCAAGGACCTCCATTTAGAAATGCTTTATTAACCTTCCTTTAGTATAGGGTAATCC 4380  
Db 4321 TGCAAGGCAAGGACCTCCATTTAGAAATGCTTTATTAACCTTCCTTTAGTATAGGGTAATCC 4380  
Qy 4381 CTTCCGGGAAACCAAGCCCGAGTCTCAGCAGAGAAACAGAAATGGGAAACCTCAGGAGG 4440  
Db 4381 CTTCCGGGAAACCAAGCCCGAGTCTCAGCAGAGAAACAGAAATGGGAAACCTCAGGAGG 4440  
Qy 4441 CAGTTTCTCCCTCGGACGGTGTAGCCACTGAAGAAAGGAAATATCTTTTCCCTGCAAC 4500  
Db 4441 CAGTTTCTCCCTCGGACGGTGTAGCCACTGAAGAAAGGAAATATCTTTTCCCTGCAAC 4500  
Qy 4501 TATCAATGGAATTTACTTAAACCTTCATCAAACTTTTCACTAGGCAATCGATAGCAC 4560  
Db 4501 TATCAATGGAATTTACTTAAACCTTCATCAAACTTTTCACTAGGCAATCGATAGCAC 4560  
Qy 4561 CCATCARATGGCCAAATCAATTTACTGGAACAGCCCTTTTCAAACTATCAAGCARAT 4620  
Db 4561 CCATCARATGGCCAAATCAATTTACTGGAACAGCCCTTTTCAAACTATCAAGCARAT 4620  
Qy 4621 AKTCAGGCGCTGTGAATGTGCAARAATAATCCCTGCTCTATCGCCAAAGCTCCCTTC 4680  
Db 4621 AKTCAGGCGCTGTGAATGTGCAARAATAATCCCTGCTCTATCGCCAAAGCTCCCTTC 4680  
Qy 4681 AGGARAACAAARAACAGGCCATTTACCTGRARAARACTGGCAACTGATTTTACCAACAG 4740  
Db 4681 AGGARAACAAARAACAGGCCATTTACCTGRARAARACTGGCAACTGATTTTACCAACAG 4740  
Qy 4741 CCCAAACCTCAGGATTTTCACTATCTACTAGTCTGGGTARATCTTTTCAAGGCTTGGGCA 4800  
Db 4741 CCCAAACCTCAGGATTTTCACTATCTACTAGTCTGGGTARATCTTTTCAAGGCTTGGGCA 4800  
Qy 4801 RAGGCGCTCCCTGTAGGACAGAAAAGGCCAAGAGGTAAATAAGGCACTAGTTTCAATGAA 4860  
Db 4801 RAGGCGCTTCCCTGTAGGACAGAAAAGGCCAAGAGGTAAATAAGGCACTAGTTTCAATGAA 4860  
Qy 4861 ATAATTTCCAGATTTCCGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAG 4920  
Db 4861 ATAATTTCCAGATTTCCGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAG 4920  
Qy 4921 GCCACAGTAAACCCAGGAGTATCCCGAGGCTTATAGGTATACGATATCACTTACCTCGGCC 4980  
Db 4921 GCCACAGTAAACCCAGGAGTATCCCGAGGCTTATAGGTATACGATATCACTTACCTCGGCC 4980  
Qy 4981 TGAAGGCCACAGTCTCAGGGAAGGTCCGAGAAAATGAATGAAYAATCAAAAGGACATCTTA 5040  
Db 4981 TGAAGGCCACAGTCTCAGGGAAGGTCCGAGAAAATGAATGAAYAATCAAAAGGACATCTTA 5040

Qy 5041 AAAAAGCAAAACCCAGGAAACCCACCTCACATGGCTGYTCTGTGTGCCCTATAGCCCTTAAAA 5100  
Db 5041 AAAAAGCAAAACCCAGGAAACCCACCTCACATGGCTGYTCTGTGTGCCCTATAGCCCTTAAAA 5100  
Qy 5101 AGAATCTGCAACTTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATCTGTATGGAAGG 5160  
Db 5101 AGAATCTGCAACTTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATCTGTATGGAAGG 5160  
Qy 5161 CCCTTCATAACCAATGACTTGTGCTTGACCAAGACAGCCAACTTAGTTGCGAGACATCA 5220  
Db 5161 CCCTTCATAACCAATGACTTGTGCTTGACCAAGACAGCCAACTTAGTTGCGAGACATCA 5220  
Qy 5221 CCTCCTTAGCCAAATATCAACAAGTTCTTAAAAATTTACAAAGAACCTTATCCCTGAGAG 5280  
Db 5221 CCTCCTTAGCCAAATATCAACAAGTTCTTAAAAATTTACAAAGAACCTTATCCCTGAGAG 5280  
Qy 5281 AGGAAAAGAACTATTTCCACCCWMTGACATGTTAGTCAAGTCCCTTCYCTCTAATT 5340  
Db 5281 AGGAAAAGAACTATTTCCACCCWMTGACATGTTAGTCAAGTCCCTTCYCTCTAATT 5340  
Qy 5341 CCCCATCCCTAGATACATCTCTGGGAAGGACCCCTACCCAGTCAATTTATTTATTTACCCCACTG 5400  
Db 5341 CCCCATCCCTAGATACATCTCTGGGAAGGACCCCTACCCAGTCAATTTATTTATTTATTTACCCCACTG 5400  
Qy 5401 CGGTTAAAGTGGCTGGAGTCTTGGATACATCACACTTGAGTCAAAATCTCTGATAC 5460  
Db 5401 CGGTTAAAGTGGCTGGAGTCTTGGATACATCACACTTGAGTCAAAATCTCTGATAC 5460  
Qy 5461 TGCCAAAGAAACCTGAAATCCAGGAGACACGCTAGCTATTTCTGTGTGAACCTCTAGAGG 5520  
Db 5461 TGCCAAAGAAACCTGAAATCCAGGAGACACGCTAGCTATTTCTGTGTGAACCTCTAGAGG 5520  
Qy 5521 ATTTGGCGCTCTCTTCAAAACAAACAGAGGAAAGTAATAAATCATATAATCTCCCC 5580  
Db 5521 ATTTGGCGCTCTCTTCAAAACAAACAGAGGAAAGTAATAAATCATATAATCTCCCC 5580  
Qy 5581 ATGSGCTCCCTTATCATATTTTCTCTKTASTGTSTTTTACCTCTTTTCTCTCTACT 5640  
Db 5581 ATGSGCTCCCTTATCATATTTTCTCTKTASTGTSTTTTACCTCTTTTCTCTCTACT 5640  
Qy 5641 GCACCCCTCTCATGCGCTGTATGACCACTAGTCTCCCTYACCMAGAGTTTCTATGGAGA 5700  
Db 5641 GCACCCCTCTCATGCGCTGTATGACCACTAGTCTCCCTYACCMAGAGTTTCTATGGAGA 5700  
Qy 5701 ATGCAAGCTCCCGGAAATATTGATGCCCATCGTATAGAGTCTTTTAAAGGGAACCCCT 5760  
Db 5701 ATGCAAGCTCCCGGAAATATTGATGCCCATCGTATAGAGTCTTTTAAAGGGAACCCCT 5760  
Qy 5761 ACCTTCAGTCCCAACCCATATGCCCGCAACTGCTATCTCTGCACTCTCTTTGCAATG 5820  
Db 5761 ACCTTCAGTCCCAACCCATATGCCCGCAACTGCTATCTCTGCACTCTCTTTGCAATG 5820  
Qy 5821 CATGCAAAATCTCATTTATTGGACAGAAAAATGATTAATCTCTGCTGAGGACTT 5880  
Db 5821 CATGCAAAATCTCATTTATTGGACAGAAAAATGATTAATCTCTGCTGAGGACTT 5880  
Qy 5881 GGAGTCACTGTCTGTGACTTATTTACCCCAAACTGGTATGCTGAATGGGGGTGGAGTT 5940  
Db 5881 GGAGTCACTGTCTGTGACTTATTTACCCCAAACTGGTATGCTGAATGGGGGTGGAGTT 5940  
Qy 5941 CAAAGTCAAGCAAGAAAGAAACATGTAAGAAAGTAATCTCCCACTCACCSSGGTACAT 6000  
Db 5941 CAAAGTCAAGCAAGAAAGAAACATGTAAGAAAGTAATCTCCCACTCACCSSGGTACAT 6000  
Qy 6001 GGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAAACCTCTCGTACC 6060  
Db 6001 GGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAAACCTCTCGTACC 6060  
Qy 6061 CATATCTCGCTGGTAAGCTATTTAATACCACTTACCTGAGGCTCATGAGGTCTCGGCC 6120  
Db 6061 CATATCTCGCTGGTAAGCTATTTAATACCACTTACCTGAGGCTCATGAGGTCTCGGCC 6120

QY 6121 CAAACCCCTACTACTGTTGGATATGCTCCCTCGAAGCTTCAGGCCATATGTTCAATC 6180  
 Db 6121 CAAACCCCTACTACTGTTGGATATGCTCCCTCGAAGCTTCAGGCCATATGTTCAATC 6180  
 QY 6181 CCTGTACCTGAACAATGGAAACAATCTCAGCACAGAAATAAACCACTTCCTCGTTTGTAGTA 6240  
 Db 6181 CCTGTACCTGAACAATGGAAACAATCTCAGCACAGAAATAAACCACTTCCTCGTTTGTAGTA 6240  
 QY 6241 GGAACCTCTGTTTCCAAATSTGGAATAAACCCATACCTCAAAACCTCACTGTGTAAATTT 6300  
 Db 6241 GGAACCTCTGTTTCCAAATSTGGAATAAACCCATACCTCAAAACCTCACTGTGTAAATTT 6300  
 QY 6301 AGCAATACATACATACAAACCACTCCCAATGATCAGGTGGGTAACTCTCCCTCCACACA 6360  
 Db 6301 AGCAATACATACATACAAACCACTCCCAATGATCAGGTGGGTAACTCTCCCTCCACACA 6360  
 QY 6361 ATAGTCTGCGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCGTTGTG 6420  
 Db 6361 ATAGTCTGCGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCGTTGTG 6420  
 QY 6421 AATGGCTCTCAGAACTATGCTCTCTCACTCTTAGTGCCCTCATGRCCTATCTAC 6480  
 Db 6421 AATGGCTCTCAGAACTATGCTCTCTCACTCTTAGTGCCCTCATGRCCTATCTAC 6480  
 QY 6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCGCAACAAGAGTACCCATT 6540  
 Db 6481 ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCGCAACAAGAGTACCCATT 6540  
 QY 6541 CTTCTCTTTTGTATAGGAGCAGGAGTGTAGTGTGCTAGTGTGCTAGTGTGCTAGTGTGCT 6600  
 Db 6541 CTTCTCTTTTGTATAGGAGCAGGAGTGTAGTGTGCTAGTGTGCTAGTGTGCTAGTGTGCT 6600  
 QY 6601 ACNACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG 6660  
 Db 6601 ACNACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG 6660  
 QY 6661 GTCCGCGACTCCTGCTGCTACCTTGCAGATCACTTAACTCTCCTAGCAGCAGTACCTTT 6720  
 Db 6661 GTCCGCGACTCCTGCTGCTACCTTGCAGATCACTTAACTCTCCTAGCAGCAGTACCTTT 6720  
 QY 6721 CRAAATCGAAGAGCTTTAGACTGCTAAGCTGTAGAGAGGGGAACTGTTTATTTT 6780  
 Db 6721 CRAAATCGAAGAGCTTTAGACTGCTAAGCTGTAGAGAGGGGAACTGTTTATTTT 6780  
 QY 6781 GGGGAAGAATGCTGTTTATTTATGTTATCAATCGGAATCGTCACTGAGAAGTTAAGAA 6840  
 Db 6781 GGGGAAGAATGCTGTTTATTTATGTTATCAATCGGAATCGTCACTGAGAAGTTAAGAA 6840  
 QY 6841 ATTCAGATCGAATACAGTACAGCAGAGAGCTTCGAAACACTGGACCTGGGCGCTC 6900  
 Db 6841 ATTCAGATCGAATACAGTACAGCAGAGAGCTTCGAAACACTGGACCTGGGCGCTC 6900  
 QY 6901 CTCAGCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6960  
 Db 6901 CTCAGCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6960  
 QY 6961 CTACTCTCTTTGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7020  
 Db 6961 CTACTCTCTTTGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7020  
 QY 7021 GAAGCTGTAAATCAAAATGAGGCCCAAGATGCGATGCAAGTACGAGTACGAGTACGAGTACGAGT 7080  
 Db 7021 GAAGCTGTAAATCAAAATGAGGCCCAAGATGCGATGCAAGTACGAGTACGAGTACGAGTACGAGT 7080  
 QY 7081 CCCCTGGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7140  
 Db 7081 CCCCTGGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7140  
 QY 7141 GAGGAATCTCAGCTGCAACAATCTCTACTACGCCCCCAATTCAGCAGGAGCAGTTAGAGC 7200  
 Db 7141 GAGGAATCTCAGCTGCAACAATCTCTACTACGCCCCCAATTCAGCAGGAGCAGTTAGAGC 7200  
 QY 7201 GGTGTGCGCCAACTCCCCAAACAGACTTATAGGTTTCTGTTGATGGGGGACTGAGA 7260

Db 7201 GGTSGTGGGCCAACCTCCCAACAGCACTTTAGGTTTTCTGTTGAGATGGGACTGAGA 7260  
 QY 7261 GACAGACTAGCTGGGATTTCTTAGGCTGATTAAGAATCCVTAAGCCTAGSTGGAGGTG 7320  
 Db 7261 GACAGACTAGCTGGGATTTCTTAGGCTGATTAAGAATCCVTAAGCCTAGSTGGAGGTG 7320  
 QY 7321 ACCACATCCACCTTTTAAACACCGGGCTTCAACTTAGYTACACCTGACCAATCAGAGAG 7380  
 Db 7321 ACCACATCCACCTTTTAAACACCGGGCTTCAACTTAGYTACACCTGACCAATCAGAGAG 7380  
 QY 7381 CTCACCTAAATGCTAATTAAGGCAACAGCAGGAGTAAAGAAATAGCAATCATYTTATTC 7440  
 Db 7381 CTCACCTAAATGCTAATTAAGGCAACAGCAGGAGTAAAGAAATAGCAATCATYTTATTC 7440  
 QY 7441 MTGAGACGACGACGAGGAGCAATGATCGGATATAAACCCCAAGTTCGAGCCGCGCAA 7500  
 Db 7441 MTGAGACGACGACGAGGAGCAATGATCGGATATAAACCCCAAGTTCGAGCCGCGCAA 7500  
 QY 7501 CGGCAACCCCTTTGGTCCCTCCCTTTGTATGGGAGCTCTGTTTCAATGCTATTTCAC 7560  
 Db 7501 CGGCAACCCCTTTGGTCCCTCCCTTTGTATGGGAGCTCTGTTTCAATGCTATTTCAC 7560  
 QY 7561 TCTATTAAATCTTGCARCTGCR 7582  
 Db 7561 TCTATTAAATCTTGCARCTGCR 7582

RESULT 3

US-09-873-367C-81  
 ; Sequence 81, Application US/09873367C  
 ; Publication No. US20030165839A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Young, Paul  
 ; APPLICANT: Soppet, Daniel  
 ; APPLICANT: Endress, Gregory  
 ; APPLICANT: Augustus, Meena  
 ; APPLICANT: Ebner, Reinhard  
 ; APPLICANT: Carter, Kenneth  
 ; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using  
 ; FILE REFERENCE: 689290-64  
 ; CURRENT APPLICATION NUMBER: US/09/873,367C  
 ; CURRENT FILING DATE: 2003-04-29  
 ; PRIOR APPLICATION NUMBER: U.S. 60/236,891  
 ; PRIOR FILING DATE: 2000-09-29  
 ; PRIOR APPLICATION NUMBER: U.S. 60/236,842  
 ; PRIOR FILING DATE: 2000-09-29  
 ; PRIOR APPLICATION NUMBER: U.S. 60/244,867  
 ; PRIOR FILING DATE: 2000-11-01  
 ; PRIOR APPLICATION NUMBER: U.S. 60/245,084  
 ; PRIOR FILING DATE: 2000-11-01  
 ; NUMBER OF SEQ ID NOS: 1067  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 81  
 ; LENGTH: 56093  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-873-367C-81

Query Match 86.1%; Score 6448; DB 3; Length 56093;  
 Best Local Similarity 95.7%; Pred. No. 0;  
 Matches 6681; Conservative 195; Mismatches 50; Indels 53; Gaps 27;  
 QY 606 GATGGAAACGTTCCCGCAAGACAAAACCGCCCTAAGAGCTATTCTGGAPAAATGGGA 665  
 Db 30953 GATGGAAACGTTCCCGCAAGACAAAACCGCCCTAAGAGCTATTCTTGGAGAAATGGGA 31012  
 QY 666 MCAATTTCCCTCAGACACTAAGAAAGAAAGCTTATTTCTTCGAGTGGCCCTG 725  
 Db 31013 CCAATTTGACCTCAGACACTAAGAAAGAAAGCTTATTTCTTCGAGTGGCCCTG 31072  
 QY 726 GCACCTCCTGAGGAAGTATAAATTATAACACCATCTTACAGCTAGACTCTTTTGTAGAA 785



QY	2945	CCAAGACTTCAGYGCARTYMTACATCGGACACTCTTGCTCTTCRGTAKGTGATGATTT	3004
DB	33266	CCAAGACTTCAGGCAATCCCTCATACCTGGACA--CTTGTCTCTTGGTAGGTGATGATTT	33323
QY	3005	ACTTTTGGYGCCTYRTTCAGAAACCTTGTGCCATCAAGCCACACCAGACRCTCTTMAATTT	3064
DB	33324	ACTTTTGGCGCCCAITTCAGAAACCTTGTGCCATCAAGCCACCCNAGGCTCTTCAATTT	33383
QY	3065	CCTCGCYACCTGTGGCTCAWGGTTTCAAACASARARGCTCARTCTGTCTCAAGCAGGT	3124
DB	33384	CCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAAAGGCTCAACTCTGTCTCAAGCAGGT	33443
QY	3125	TAAATACTTAGGCTTAARATTTATCCAAAGGCAACAGGCCCTCAGTGAGGAAATRYATCCA	3184
DB	33444	----TACTTAGGCTTAARATTTATCCAAAGGCAACAGGCCCTCAGTGAGGAAACACATCCA	33499
QY	3185	GCCTATACCTGGCTTATCCTCATCYCAAAACCTTAAAGCAACTAAGRRRTTCTTGGCRT	3244
DB	33500	GCCTATACCTGGCTTATCCTCATCCCAAAACCTTAAAGCAACTAAGGGGATCTTGGCGT	33559
QY	3245	AYAGGYTTCGCGGAATFGAATCCCGAGGTWGGCRAAATAGCCAGGYCATTAAWATA	3304
DB	33560	AATAGGTTTCTGCGGAAATGGATT--CCAGGTATGGCGAAATAGCCAGGTCAATTAATA	33618
QY	3305	CASRAATTAGGAACTCAGAAAGCCATACCCATTARTTAAGATGAYAMCTGAAGYMR	3364
DB	33619	CACATAATTAGGAACTCAGAAAGCCATACCCATTARTTAAGATGAGCAACTGAAAGTAG	33678
QY	3365	AAGTGGCTTTCAGGCGCCCTTAAAGAAGGCCTTAAACCCCAAGYCCCACTGTTAAGYTTGCC	3424
DB	33679	AAGTGGCTTTCAGGCGCC-----TAAACCCCAAGGCCCACTGTTAAGYTTGCC	33724
QY	3425	AACRGGGCAAGACTTTTSTYATAYRTCAAGAAAAAACAAGRAAYAGCTCTRGAGTCC	3484
DB	33725	AACAGGCAAGACTTTTCTTCAATGTCAAG--AAAAACAGGAATAGCTCTAGGAGTCC	33783
QY	3485	TTACACAGTCCBAGGAGAGCTTCGAACCTRTGCRVACCTGASTAAGGAAYTGATG	3544
DB	33784	TTACACAGATCCGAGGAGTGGCTTGAACCTGTGGCATACCTGCACTAAGGAATATGATG	33843
QY	3545	TAGTGGCAAAAGGTTGRCYTCATTGTTTAYGGTAGTGGTGGCAGTAGCAGTGTAGTAT	3604
DB	33844	TAGTGGCAAAAGGTTGACCTCATGTTTACGGGTAGTGGTGGCAGTAGCAGTCTTAGTAT	33903
QY	3605	CTGAAGCAGTTAAATAATAACAGGGAAGATCTTACTGTGTGGACATCTCATGAGTGA	3664
DB	33904	CTGAAGCAGTTAAATAATAACAGGGAAGATCTTACTGTGTGGACATCTCATGATGTGA	33963
QY	3665	AYRGATATCTACTGTAAAGGAGACTTGTGGCTGTGCAGCAACYGTGTTTACTTAATRTC	3724
DB	33964	ATGGCATACTCACTGTCTAAGGAGACTTGTGGCTGTGCAGCAACTGTTTACTTAATGTC	34023
QY	3725	AGGCTCTATTACTTGAAGGCGCAGTCTGCRACCTGTGCACTGTGCACTCTTAACCCAG	3784
DB	34024	AGGCTCTATTACTTGAAGGCGCAGTCTGCGACTGTGCACTGTGCACTCTTAACCCAG	34083
QY	3785	YCNCAFTTCTTCAGACAAATGAAGAAAGATARAAYATAACTGTCAACAARTAAATTCCTC	3844
DB	34084	CCACATTTCTTCAGACAAATGAAGAAAGATARAAYATAACTGTCAACAARTAAATTCCTC	34143
QY	3845	AAACCTATGCCACTCGAGGGGACCTTGTAGRTTCCTTGTGATCTGATCCYGCACCTTCAAC	3904
DB	34144	AAACCTATGCCACTCGAGGGGACCTTGTAGRTTCCTTGTGATCTGATCCCGACC--TCAC	34202
QY	3905	TTGTATCTAGTGGAGTTCCTTGTAGAAAGGACCTTCGAAAGYGGGGTATGCAGTG	3964
DB	34203	TTGTATCTAGTGGAGTTCCTTGTAGAAAGGACCTTCGAAAGYGGGGTATGCAGTG	34262
QY	3965	GTCAAGTAAATGGGAATAYTTGAAAGTAAATCCCTCCTCAGTCCAGGAACCTAGTCTAGCTR	4024
DB	34263	GTCAAGTAAATGGGAATAYTTGAAAGTAAATCCCTCCTCAGTCCAGGAACCTAGTCTAGCTA	34322
QY	4025	GCAGAACTAATAGCCVTCAYTTGGGCACTAGAAATTAGGAGAAAGAAAGGGYAAATATA	4084
DB	34323	GCAGAACTAATAGCCCTCACTTGGGCACTAGAAATTAGGAGAAAGAAAGGGYAAATATA	34382
QY	4085	TATACAGACTCTTARTATCTYACCTAGTCNTCCATGCCCATCMRGCAATATGSARAGAA	4144
DB	34383	TATACAGACTCTTARTATCTYACCTAGTCNTCCATGCCCATCMRGCAATATGSARAGAA	34442
QY	4145	AGGGAATTCCTAACTTCYAGRGGAACACCTTATCAMACATCAGGAAGCCATTAGBARATTA	4204
DB	34443	AGGGAATTCCTAACTTCYAGRGGAACACCTTATCAMACATCAGGAAGCCATTAGBARATTA	34502
QY	4205	TTATYTCGCTGACAGAAACCTTAPAGAGGTGGWAGTCTTACACTGTCTGGGGTCTATCANAAA	4264
DB	34503	TTATYTCGCTGACAGAAACCTTAPAGAGGTGGWAGTCTTACACTGTCTGGGGTCTATCANAAA	34562
QY	4265	GGAAAGRAAAGGGAATASAGRGAAATGCCAAGCAKATATTCGAAGCAAAAGAGCTGCA	4324
DB	34563	GGAAAGRAAAGGGAATASAGRGAAATGCCAAGCAKATATTCGAAGCAAAAGAGCTGCA	34622
QY	4325	AGGCAAGACCTTCCATTAGAAATGCTTATTAAACCTTCCCTTCTAGTATAGGCTAATCCTTTC	4384
DB	34623	AGGCAAGACCTTCCATTAGAAATGCTTATTAAACCTTCCCTTCTAGTATAGGCTAATCCTTTC	34682
QY	4385	CGGGAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCAGAGG--CAG	4443
DB	34683	CGGGAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCAGAGGAGCAG	34742
QY	4444	TTTTCTCCCTCGGAGCGGTAGCCACTGAAGAGGGAATATCTTTTGGCTGCACTAT	4503
DB	34743	TTTTCTCCCTCGGAGCGGTAGCCACTGAAGAGGGAATATCTTTTGGCTGCACTAT	34802
QY	4504	CCAATGGAAATTAATTAACCCCTTCAATCAAAACCTTCTACTTAGGCATCGATAGCACCCA	4563
DB	34803	CCAATGGAAATTAATTAACCCCTTCAATCAAAACCTTCTACTTAGGCATCGATAGCACCCA	34862
QY	4564	TCARATGGCAAAATCATTTATTTATGGAACAGGCTTTTCAAAACTATCAAGCARATAKT	4623
DB	34863	TCARATGGCAAAATCATTTATTTACTGGACAGGCTTTTCAAAACTATCAAGCAGATAGT	34922
QY	4624	CAGGCGCTGTGAATGTGCCABARAATAATCCCTGCTCTATCGCAAGCTCCTTCAGG	4683
DB	34923	CAGGCGCTGTGAATGTGCCABARAATAATCCCTGCTCTATCGCAAGCTCCTTCAGG	34982
QY	4684	ARAAACAARAAACAGGCCATTACCTGRRARAARACTGGCAACTGATTTTCCCAACAGCCC	4743
DB	34983	ARAAACAARAAACAGGCCATTACCTGRRARAARACTGGCAACTGATTTTCCCAACAGCCC	35042
QY	4744	AAACCTCAGGGATTTTCACTACTACTAGTCTGGGTARATPACTTTTCAAGGCTTGGGCGARAG	4803
DB	35043	AAACCTCAGGGATTTTCACTACTACTAGTCTGGGTARATPACTTTTCAAGGCTTGGGCGARAG	35102
QY	4804	GCCTTCCCTGTAGACAGAAAGGCCCAAGAGGTAAATAAAGGCATAGTTCATGAATAA	4863
DB	35103	GCCTTCCCTGTAGACAGAAAGGCCCAAGAGGTAAATAAAGGCATAGTTCATGAATAA	35162
QY	4864	ATTCCAGATTCGGAGCTTCCCGAGGCTTACAGAGTGAACAATAGCCCTCTTCCAGGCC	4923
DB	35163	ATTCCAGATTCGGAGCTTCCCGAGGCTTACAGAGTGAACAATAGCCCTCTTCCAGGCC	35222
QY	4924	ACAGTAACCCAGGGAGTATCCAGGCGTTTAGGTATACGATATCACTTACCTGCGCTGA	4983
DB	35223	ACAGTAACCCAGGGAGTATCCAGGCGTTTAGGTATACGATATCACTTACCTGCGCTGA	35282
QY	4984	AGGCCACAGTCTCAGGGAAGGTTCAGAAATGAATGAATAAAYACTCAAGGACATCTTAAAA	5043
DB	35283	AGGCCACAGTCTCAGGGAAGGTTCAGAAATGAATGAATAAAYACTCAAGGACATCTTAAAA	35342
QY	5044	AAGCAAAACCCAGGAAACCCACTCAGTGGCTGTCTGTTGCTTATAGCCTTAAAGAGA	5103
DB	35343	AAGCAAAACCCAGGAAACCCACTCAGTGGCTGTCTGTTGCTTATAGCCTTAAAGAGA	35402
QY	5104	ATCTGCAACTTTTCCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCC	5163



|||||  
35403 ATCTGCACTTTCCCAAAAGACGAGCTTAGCCCATAGAAATGCTGTATGGAGGGCC 35462  
|||||  
5164 TTCTAAACCAATGACTTGTGCTTGACCCAGACAGCCAACTTAGTTCGACAGATCACT 5223  
|||||  
35463 TTCTAAACCAATGACTTGTGCTTGACCCAGACAGCCAACTTAGTTCGACAGATCACT 35522  
|||||  
5224 CCTTAGCCCAATATCAACAAGTTCCTTAAACATTTAACAAGTTCATCCCTGAGNAGAG 5283  
|||||  
35523 CCTTAGCCCAATATCAACAAGTTCCTTAAACATTTAACAAGTTCATCCCTGAGNAGAG 35582  
|||||  
5284 GAAAGAACTATTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCVCTCTAATCCC 5343  
|||||  
35583 GAAAGAACTATTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCVCTCTAATCCC 35642  
|||||  
5344 CATCCCTAGATACATPCTCTGGGAAGACCTTACCCAGTCAATTTATATACCCCACTGCGG 5403  
|||||  
35643 CATCCCTAGATACATPCTCTGGGAAGACCTTACCCAGTCAATTTATCTACCCCACTGCGG 35702  
|||||  
5404 TTAAAGTGGCTGGAGTGGAGTCTTGATACATCACTTGAGTCAAACTCTGGATCTGC 5463  
|||||  
35703 TTAAAGTGGCTGGAGTGGAGTCTTGATACATCACTTGAGTCAAACTCTGGATCTGC 35762  
|||||  
5464 CAAAGAACTGAAATCCAGAGACAAACGCTAGCTATTCTCTGTGAACCTCTAGAGGATT 5523  
|||||  
35763 CAAAGAACTGAAATCCAGAGACAAACGCTAGCTATTCTCTGTGGAACCTCTAGAGGATT 35822  
|||||  
5524 TCGCCCTGCTCTTCAAAACAAACAGGAGGAAAGTAACTAAATCATAAATCCCCCATG 5583  
|||||  
35823 TCGCCCTGCTCTTCAAAACAAACAGGAGGAAAGTAACTAAATCATAAAT - CCCCATG 35881  
|||||  
5584 GSCCTCCCTTATCATATTTCTCTKTAAGTACAGTCTCTTACCTCTTACCTCTACTGCA 5643  
|||||  
35882 GSCCTCCCTTATCATATTTCTTACTGTCTTTTACCTCTTACCTCTTCACTCTCACTGCA 35941  
|||||  
5644 CCCCCTCCATGCGGTGTATGACCACTAGCTCCCTCTACCMAGATTCTATGAGAAATG 5703  
|||||  
35942 CCCCCTCCATGCGGTGTATGACCACTAGCTCCCTCTACCAAGAGTTCTATGAGAAATG 36001  
|||||  
5704 CAGCGTCCGGAAATATATGATGCCCATCGTATAGGAGTCTTTSTAAGGAAACCCCCACC 5763  
|||||  
36002 CAGCGTCCGGAAATATATGATGCCCATCGTATAGGAGTCTTTSTAAGGAAACCCCCACC 36061  
|||||  
5764 TTCACCTGCCCAACCCCATATGCCCAGCAACCTGCTATCACTCTGCCACTCTTTGATGAT 5823  
|||||  
36062 TTCACCTGCCCAACCCCATATGCCCAGCAACCTGCTATCACTCTGCCACTCTTTGCAATG 36121  
|||||  
5824 GCAAACTACTCATTTTGGACAGGAAATATGATTAATCCTAGTGTCTTGAGGACTTGA 5883  
|||||  
36122 GCAAACTACTCATTTTGGACAGGAAATATGATTAATCCTAGTGTCTTGAGGACTTGA 36181  
|||||  
5884 GTCACCTGTCTGTGGACTTACTTACCCAAACTGGTATGCTGTATGCTGGGGTGGAGTTCAA 5943  
|||||  
36182 GTCACCTGTCTGTGGACTTACTTACCCAAACTGGTATGCTGTATGCTGGGGTGGAGTTCAA 36241  
|||||  
5944 GATCAGGCAAGGAAACATGTATAAGAAATATCTCCCACTCAACGCGGTAATGGC 6003  
|||||  
36242 GATCAGGCAAGGAAACATGTATAAGAAATATCTCCCACTCAACGCGGTAATGGC 36301  
|||||  
6004 ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCAT 6063  
|||||  
36302 ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACCCAT 36361  
|||||  
6064 ACTCGCTGGTAAGCCTATTTAATACCACTTACCTGCGGCTCCATGAGTCTCGGCCCAA 6123  
|||||  
36362 ACTCGCTGGTAAGCCTATTTAATACCACTTACCTGCGGCTCCATGAGGTCTCGGCCAA 36421  
|||||  
6124 AACCCCTAATACTGTGTGGATATGCTCCCTGGAACCTTCAGGCAATATGTTTCAATCCCT 6183  
|||||  
36422 AACCCCTAATACTGTGTGGATATGCTCCCTGGAACCTTCAGGCAATATGTTTCAATCCCT 36481  
|||||  
6184 GTACCTGAAACATGGAACAACTTTCAGCACAGAAATAAACCACTTCCGTTTTAGTAGGA 6243  
|||||

Db  
36482 GTACTGTAAACAATGGAAACAACCTTCAGCACAGAAATAAACACCACTTCCTCGTTTTAGTAGGA 36541  
Qy  
6244 CCTCTTGTTTCCAATSTGGAAATAAACCATACCTCAAACTCACTGTGTGTAAATTTTAGC 6303  
Db  
36542 CCTCTTGTTTCCAATSTGGAAATAAACCATACCTCAAACTCACTGTGTGTAAATTTTAGC 36601  
Qy  
6304 AATACTACATACACAACAACCTCCCAATGCACTAGGTGGGTAACTCTCTCCCAACAATA 6363  
Db  
36602 AATACTACATACACAACAACCTCCCAATGCACTAGGTGGGTAACTCTCTCCCAACAATA 36661  
Qy  
6364 GTCTGCCCTACCTCAGGAATAATTTTTGTCTGTGGTACCTCAGGCTATTCGTTGTTTGAAT 6423  
Db  
36662 GTCTGCCCTACCTCAGGAATAATTTTTGTCTGTGGTACCTCAGGCTATTCGTTGTTTGAAT 36721  
Qy  
6424 GGCTCTCAGAAATCTATGTGCTTCTCTCATCTTCTAGTGCCCCCACTATGRCATCTACACT 6483  
Db  
36722 GGCTCTCAGAAATCTATGTGCTTCTCTCATCTTCTTAGTGCCCTTATGACCATCTACACT 36781  
Qy  
6484 GAAACAAGATTTATACAGTTATGTCTATCTAAGCCCGCAACAAAGAGTAGTACCCATTTCTT 6543  
Db  
36782 GAACAAGATTTATACAGTTATGTCTATCTAAGCCCGCAACAAAGAGTAGTACCCATTTCTT 36841  
Qy  
6544 CCTTTGTTATAGGACGAGGAGTCTAGGTGCACTAGGTACTGGCAATTTGGCGGTATCAACA 6603  
Db  
36842 CCTTTGTTATAGGACGAGGAGTCTAGGTGCACTAGGTACTGGCAATTTGGCGGTATCAACA 36901  
Qy  
6604 ACCTCTACTAGTTCTACTACAACTATCTCAAGAACTTAAATGGGGACATGGAAAGGCTC 6663  
Db  
36902 ACCTCTACTAGTTCTACTACAACTATCTCAAGAACTTAAATGGGGACATGGAAAGGCTC 36961  
Qy  
6664 GCCGACTCCCTGGTCACTTCGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTCTCRA 6723  
Db  
36962 GCCGACTCCCTGGTCACTTCGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTCTCRA 37021  
Qy  
6724 AATCGAAGAGCTTTAGACTTCTAACCTCGARAGAGGGGAACTGTTTATTTTTAGGG 6783  
Db  
37022 AATCGAAGAGCTTTAGACTTCTAACCTCGARAGAGGGGAACTGTTTATTTTTAGGG 37081  
Qy  
6784 GAAAGATGCTGTTATTTATTTAATCAATCCGGAATTCGTCACCTAGAAAGTTTAAAGAAAT 6843  
Db  
37082 GAAAGATGCTGTTATTTATTTAATCAATCCGGAATTCGTCACCTAGAAAGTTTAAAGAAAT 37141  
Qy  
6844 CSAGATCGAATACAACTGAKAGAGAGCTTCGAAACACTTGAACCTCGGGGCTCTCTC 6903  
Db  
37142 CGAGATCGAATACAACTGAKAGAGAGCTTCGAAACACTTGAACCTCGGGGCTCTCTC 37201  
Qy  
6904 AGCCRAATGGATGCCCTGGATTCTCCCTTTCTTAGGACCTCTAGCAGCTATATATTTGCTA 6963  
Db  
37202 AGCCRAATGGATGCCCTGGATTCTCCCTTTCTTAGGACCTCTAGCAGCTATATATTTGCTA 37261  
Qy  
6964 CTCTCTTTGGACCCCTGTATCTTTTACCTCTTGTAACTTTGTCTCTTCCAGAAATCGAA 7023  
Db  
37262 CTCTCTTTGGACCCCTGTATCTTTAACTCTTGTAACTTTGTCTCTTCCAGAAATCGAA 37321  
Qy  
7024 GCTGTAAACTTACAAATGGAGCCCAAGATGAGTCCAAGACTTAAGATCTACCCGAGACCC 7083  
Db  
37322 GCTGTAAACTTACAAATGGAGCCCAAGATGAGTCCAAGACTTAAGATCTACCCGAGACCC 37381  
Qy  
7084 CTGACCCGGCTGTAGCCCACTGATCTGATTTTAAATGACATCAAAAGGACCCCTCTCTGAG 7143  
Db  
37382 CTGACCCGGCTGTAGCCCACTGATCTGATTTAATGACATCAAAAGGACCCCTCTCTGAG 37441  
Qy  
7144 GAAATCTCAGCTGCAAACTCTCTACTAGCCCAATTTAGCGGAGGAGCGATTAGAGCGGT 7203  
Db  
37442 GAAATCTCAGCTGCAAACTCTCTACTAGCCCAATTTAGCGGAGGAGCGATTAGAGCGGT 37501  
Qy  
7204 SSTCGGCCAACTCCCAACAGCACTTAGGTTTTCTGTTGATATGGGGGACTTGAGAGAC 7263  
Db  
37502 CTTGCGGCCAACTCCCAACAGCACTTAGGTTTTCTGTTGATATGGGGGACTTGAGAGAC 37561  
Qy  
7264 AGGACTAGCTGGATTTCTTAGCTGATTAAGAAATCCYTAAGCCTTAGSTGGGAGGTGACC 7323  
Db  
37562 AGGACTAGCTGATTTCTTAGCTGATTAAGAAATCCCTAAGCCTTAGCTGGGAGGTGACC 37621



```
QY 7324 ACATCCACCTTTAAACGCGGCTTGCAACTTAGTTCACACCTGACCAATCAGAGAGCTC 7383
Db 37622 ACATCCACCTTTAAACGCGGCTTGCAACTTAGTTCACACCTGACCAATCAGAGAGCTC 37681
QY 7384 ACTAAATGCTAAATTAGGCAAGACAGGAGGTAAAGAAATAGCCAAATCATYTTATTGCMTG 7443
Db 37682 ACTAAATGCTAAATTAGGCAAGACAGGAGGTAAAGAAATAGCCAAATCATYTTATTGCGTG 37741
QY 7444 AGAGCAGCAGGAGGAGCAATGATCGGGATATAAACCCAAAGYTTTCGAGCCGCGCAACGG 7503
Db 37742 AGAGCAGCAGGAGGAGCAATGATCGGGATATAAACCCAAAGYTTTCGAGCCGCGCAACGG 37801
QY 7504 CAACCCCTTTGGTCCCTCCCTTTCTGATGGGAGCTCTGTTTTCATGCTATTTCACCTCT 7563
Db 37802 CAACCCCTTTGGTCCCTCCCTTTCTGATGGGAGCTCTGTTTTCATGCTATTTCACCTCT 37861
QY 7564 ATTAATCTTTGCARCTGCR 7582
Db 37862 ATTAATCTTTGCARCTGCA 37880

RESULT 4
US-10-843-641A-81
; Sequence 81, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 81
; LENGTH: 56093
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-81

Query Match 86.1%; Score 6448; DB 10; Length 56093;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 6681; Conservative 195; Mismatches 50; Indels 53; Gaps 27;

QY 606 GATGGAAAGCTTCCCGCAAGCAAAACGCCCCCTAAGACGTAATTCGGAAATTTGGGA 665
Db 30953 GATGGAAAGCTTCCCGCAAGCAAAACGCCCCCTAAGACGTAATTCGGAAATTTGGGA 31012
QY 666 MCAATTTGACCTTCAGACACTAGAAAGAACGACTTATATCTTCTGCGAGTCCGCTG 725
Db 31013 CCAATTTGACCTTCAGACACTAGAAAGAACGACTTATATCTTCTGCGAGTCCGCTG 31072
```

```
QY 726 GCACCTCTGAGGGAAGTATATAATTAACACCATCTTTACAGTAGAGCYTCTTTTGTAGAA 785
Db 31073 GCACCTCTGAGGGAAGTATATAATTAACACCATCTTTACAGTAGAGCYTCTTTTGTAGAA 31132
QY 786 AAGGCAAAATGGAGTGAAGTGCATTAAGTACAAACTTTCTTTTCAATTAAGAGACAACCTCAC 845
Db 31133 AAGGCAAAATGGAGTGAAGTGCATTAAGTACAAACTTTCTTTTCAATTAAGAGACAACCTCAC 31192
QY 846 AATTATGTAAAAAGTGTGATTTATGCCCCA CAGGAAGCCTTTCAGAGTCTACCTCCCTATC 905
Db 31193 AATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTTCAGAGTCTACCTCCCTATC 31252
QY 906 CGAGATCCCGGACTCTCTCCCCCAATTAAGGAACCCCTCTCAACCCAAATGGTCCAA 965
Db 31253 CGAGATCCCGGACTCTCTCCCCCAATTAAGGAACCCCTCTCAACCCAAATGGTCCAA 31312
QY 966 AAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCTAATTTCCCAATTTATGA 1025
Db 31313 AAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCTAATTTCCCAATTTATGA 31372
QY 1026 CCCCTCCCAAGCAGTGGGAGGAGAGAAATTCGGCCAGCAGAGTGCATGTGCVTTTTY 1085
Db 31373 CCCCT-CCAGCAGTGGGAGGAGAGAAATTCGGCCAGCAGAGTGCATGTGCTTTTTC 31431
QY 1086 TCTCCAGACTTAAAGCAAAATAAAAA CAGACTTTAGGTAAATTTCTCAGATAAYCCTGATGG 1145
Db 31432 TCTCCAGACTTAAAGCAAAATAAAAA CAGACTTTAGGTAAATTTCTCAGATAAYCCTGATGG 31491
QY 1146 CTATATTGTGTTTTACAGGGTTAGACAAATTTCTTGATCTCAGATGGAGAGATATATA 1205
Db 31492 CTATATTGTGTTTTCAAGGGTTAGACAAATTTCTTGATCTCAGATGGAGAGATATA-A 31550
QY 1206 TGTCACTGTCTAAATCAGACACTAAACCCCAATCAGAGAGTGCACCATAACTGCAGCCT 1265
Db 31551 TGTCACTGTCTAAATCAGACACTAAACCCCAATCAGAGAGTGCACCATAACTGCAGCCT 31610
QY 1266 GAGGTTTTGGCGATCTCTGGTATCTCAGTCAGTCAATGGATANGGATGACAA CAGAAAG 1325
Db 31611 GAGGTTTTGGCGATCTCTGGTATCTCAGTCAGTCAATGG-ATAGGATGACAA CAG-AGG 31667
QY 1326 AAGANAATGATTTCCCA CAGGCGAGCGAGTTCCTCAGTCTASACCTCATTTGGGAC 1385
Db 31668 AAGAGAATGATTTCCCA CAGGCGAGCGAGTTCCTCAGTCTAGACCTCATTTGGG- 31724
QY 1386 ACAGAAATCAGTAAACATGGGAGATTTGGTCTG CAGACATTTGCTAATCTTGCTGTGCTASAA 1445
Db 31725 ACAGAAATCAGAAAT-GGAGTTGGTCTG CAGACATTTGCTAATCTTGCTGTGCTAGAA 31783
QY 1446 GGACTAAGGAAAACTASGAAGAAATCTAYGAATTA CTAATGATGTCACCACTAACACA 1505
Db 31784 GGACTAAGGAAAACTAGGAAG-AGTCTATGAATTA CTAATGATGTCACCACTAACACA 31842
QY 1506 GGGGAAGGGAAGAAATCTCTGCTTTCTGAGAGAGACTAAGGGAGGCAATGAGGAAGC 1565
Db 31843 -GGGAAGGGAAGAAATCTCTGCTTTCTGAGAGAGACTAAGGGAGGCAATGAGGAAGC 31901
QY 1566 GTGCTCTCTGTCACTGACTCTCTGAGGCGCACTAATCTTAAAGCGTAAGTTTATCA 1625
Db 31902 GTGCTCTCTGTCACTGACTCTCTGAGGCGCACTAATCTTAAAGCGTAAGTTTATCA 31961
QY 1626 CTCAGTCAGCTGCAGACATTAG-AAAAAATTTCAAAGTCTGCGTAGGCCCGGAGCAAA 1684
Db 31962 CTCAGTCAGCTGCAGACATTAGAAAAAATTTCAAAGTCTGCGTAGGCCCGGAGCAAA 32021
QY 1685 ACTTAGAAA CTTATTTGAACCTTTGGCAACCTCGGTTTTTATATAGAGATCAGGAGGAGC 1744
Db 32022 ACTTAGAAA CTTATTTGAACCTTTGGCAACCTCGGTTTTTATATAGAGATCAGGAGGAGC 32081
QY 1745 AGCGGACAGGACAAA CCGGATTAATAAAAAAGGCAACCGCTTTTAGTCTATGACCTCAGG 1804
Db 32082 AGCGGACAGGACAAA CCGGATTAATAAAAAAGGCAACCGCTTTTAGTCTATGACCTCAGG 32141
```

1805 CAAGTGGAGCTTTGGAGGCTCTGGAAGAGGAAAGCTGGGCAAAATTGAATGCTAATAGG 1864  
Db  
32142 CAAGTGGAGCTTTGGAGGCTCTGGAAGAGGAAAGCTGGGCAAAATTGAATGCTAATAGG 32201  
Qy  
1865 GCTTGTCTTCAGTGGGGTCTCAAGGACACTTTTAAAAAGATGTGCCAAGTAGAAGTAAG 1924  
Db  
32202 GCTTGTCTTCAGTGGGGTCTCAAGGACACTTTTAAAAAGATGTGCCAAGTAGAAGTAAG 32261  
Qy  
1925 CGCGCCCTTGTGTCATGCCCCCTTATTTCAAGGAATCACTGGAAGGCCCACTGCCCCAGG 1984  
Db  
32262 CGCGCCCTTGTGTCATGCCCCCTTATTTCAAGGAATCACTGGAAGGCCCACTGCCCCAGG 32321  
Qy  
1985 GGCAAAAGGTCTTTTCAAGTCAAGGACCACTAACACAGATGATCCAGCAGGACTGAGGG 2044  
Db  
32322 GGCAAAAGGTCTCTGAGTCAGAGCCACTAACACAGATGATCCAGCAGGACTGAGGG 32381  
Qy  
2045 TGCTGGGGCAAGCGCCATCCCATGCCATCACCTCACAGAGCCCTGGGTATGCTTGACC 2104  
Db  
32382 TGCTGGGGCAAGCGCCATCCCATGCCATCACCTCACAGAGCCCTGGGTATGCTTGACC 32441  
Qy  
2105 ATGAGGGCCAGGAAGTGTCTCTGGGACA CTGTGTGCGGTCTTCTTAGTCTTATCTCTTC 2164  
Db  
32442 ATGAGGGCCAGG-AGGTGTCTCTCTGGGACA CTGTGTGCGGTCTTCTTAGTCTTATCTCTTC 32500  
Qy  
2165 TGTCCGGGCAACTGTCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAGAGCGGCA 2224  
Db  
32501 TGTCCGGGCAACTGTCTCCAGATCTGTCACTA-TCCTGAGGGGGTCC-TAAGACGGGCA 32558  
Qy  
2225 GTCACTAGATACTTTTCCAGCCACTAAGATTATGAATCGGGGAGCTTTATTTCTTTTCA 2284  
Db  
32559 GTCACTAGATAC-TTCTCCAGCCACTAAGATTATG-ACTGGGGGAGCTTTATTTCTTTTNC 32616  
Qy  
2285 ATGCTTTTCTAATATTATGCTTGAAGCCCACTACTCTTGTGTAAGGAGAGACATTCAGCAA 2344  
Db  
32617 ATGCTTTTCTAATATTATGCTTGAAGCCCACTACTCTTGTGTAAGGAGAGACATTCAGCAA 32676  
Qy  
2345 RAGCAGGGGCCATATACACCTGACATAGGAGGACACCCGTTTGTCTGTCNCCCTG 2404  
Db  
32677 AAGCAGGGGCCATATATACCTGACATAGGAGGAGAACACCCGTTTGTGTG-CCCTG 32735  
Qy  
2405 CTTGAGGAAGGAATTAATCCTGAAAGTCTGGGCAACAGAGGACAAATATGAGCAGGCCAAA 2464  
Db  
32736 CTTGAGGAAGGAATTAATCCTGAAAGTCTGGGCAACAGAGGACAAATATGAGCAG-CAAA 32794  
Qy  
2465 GAATGCCCTCTCTGTTCAAGTTAACTAAAGGATTCACCTCTTCCCTCACAAAGGCA 2524  
Db  
32795 GAATGCCCTCTCTGTTCAAGTTAACTAAAGGATTCACCTCTTCCCTCACAAAGGCA 32854  
Qy  
2525 GTACCCCTCAGACCAAGGCCCAAGGATTCCAAAGATTTGTTAAGGACTTAAAGC 2584  
Db  
32855 GTACCCCTCAGACCAAGGCCCAAGGACTTCCAAAGATTTGTTAAGGACTTAAAGC 32914  
Qy  
2585 CCAAGGCTTAGTAAAAACCATGCACTACTCCCTGCAAGTAACTCCGTAGTGGATTGAGGAGG 2644  
Db  
32915 CCAAGGCTTAGTAAAAACCATGCAAGTAACTCCCTGCAAGTAACTCC-AAAAATTAGGAG 32967  
Qy  
2645 CACAGAAACCAAGTGGACAGTGGAGGGTTAGTGCAGAGATCTCAGGATTAATCAATGAGGC 2704  
Db  
32968 TACAGAAACCAACAGACAGTGG-AGTTAGTGCAGAGATCTCAGGATTAATCAAT-AGGCG 33025  
Qy  
2705 CGTTGTCTTTATACCCAGCTGTACTAGCCCTTATACGTGTGMYTTCCAAATACCAGA 2764  
Db  
33026 TGTGTGTCTTATACCCAGCTGTACTAGCCCTTATACGTGTGMYTTCCAAATACCAGA 33085  
Qy  
2765 GGAAGCAGAGTGGTTTACACSTCTCTGACCTTMAAGATGCTCTTCTGCACTCCCTGTACA 2824  
Db  
33086 GGAAGCAGAGTGGTTTACAGTCTCTGACCTTCAAGGATGCTCTTCTGCACTCCCTGTACA 33145  
Qy  
2825 TCTGACTCTCAATCTTGTGTGCTTTGAAGATCTTCAAAACCAACATCTCAACTCAC 2884  
Db  
33146 TCTGACTCTCAATCTTGTGTGCTTTGAAGATCTTCAAAACCAACATCTCAACTCAC 33205  
Qy  
2885 CTGACTRTTTTACCCCAAGGGTTACGGGATAGYCCCCATCTATTTTGGCCAGGCAATTAGC 2944

33206 CTGACATATTTTACCCCAAGGTTTCAGGGATAGTCCCATCTATTGTCAGGCACTTAGC 33265  
Qy  
2945 CCAAGACTTGGAGYCAATYMTTCATCTGGGACACTCTTGTCTCTTCTGTAKGTGATGATTTT 3004  
Db  
33266 CCAAGACTTGGAGCAATCTCATCTCTGGACA--CTTGTCTCTTGGTAGGTGATGATTTT 33323  
Qy  
3005 ACTTTTRGCGCCYRWTTCAGAAACCTTGTGCCATCAAGCCCAAGCCTCTCTTAAATTT 3064  
Db  
33324 ACTTTTGGCCGCCCATTCAGAAACCTTGTGCCATCAAGCCCAAGCGCTCTTCAATTT 33383  
Qy  
3065 CTTGCYACCTGTGTGGCTACAWGGTTTCCAAACSPARARGCTCACTCTGTCTCACAGCAGGT 3124  
Db  
33384 CTTGCTACCTGTGTGGCTACATGGTTTCCAAACCAAGGCTCACTCTGCTCACAGCAGGT 33443  
Qy  
3125 TAAATACCTTAGGRCTTAABATTATCCAAAGGCCACCAAGGCCCTCAGTAGGAAATATCA 3184  
Db  
33444 ----TACTTGGGCTTAAATTTATCCAAAGGCCACCAAGGCCCTCAGTAGGAAACATCCA 33499  
Qy  
3185 GCCTATACCTGGCTTATCTCATCTCYCAAAACCTTAAGCAACTAAGGRRTTCTTTGGCT 3244  
Db  
33500 GCCTATACCTGGCTTATCTCATCTCCAAACCTTAAAGCAACTAAGGGGATTCCTTTGGCT 33559  
Qy  
3245 AAYAGGYTTCTGCCGAATGATTTCCCGAGTWTGGCRAAATAGCCAGGYCATTAATA 3304  
Db  
33560 AATAGGTTTCTGCCGAATGATTT-CCAGGATGTGGCAATAGCCAGGTCAATTAATA 33618  
Qy  
3305 CASTAATTAAGGAACTCAGAAAGCAATACCCATTARTTAAGTAGGAYAMCTGAAGYMR 3364  
Db  
33619 CACTAATTAAGGAACTCAGAAAGCAATACCCATTARTTAAGTAGGAYAMCTGAAGYMR 33678  
Qy  
3365 AAGTGGCTTTCCAGGCCCTTAAAGAGGCTTAAACCCAGCYCCAGGTGTTAAGYTTGCC 3424  
Db  
33679 AAGTGGCTTTCCAGGCC-----TAAACCCAGGCCAGGTGTTAAGYTTGCC 33724  
Qy  
3425 AACGGGCAAGACTTTTSTTATAYRTTCAAGAAAAAACAAGAAAYAGCTCTGGAGTCC 3484  
Db  
33725 AACAGGCAAGACTTTTCTCATATGTACAG-AAAAAACAGGAATAGCTCTAGGAGTCC 33783  
Qy  
3485 TTACACAGTCCRAGGGAYAGCTTGGCAACCVRTGGYACCTGASTAAGGAAAYTGATG 3544  
Db  
33784 TTACACAGTCCRAGGGATGAGCTTGGCAACCTGTGGCATACCTGACTAAGGAAATGATG 33843  
Qy  
3545 TAGTGGCAAGGGTTGRCYCTCATTTATYGGGTAGTGGTGGCAGCTAGCAGYTKTAGTAT 3604  
Db  
33844 TAGTGGCAAGGGTTGRCYCTCATTTATYGGGTAGTGGTGGCAGCTAGCAGYTKTAGTAT 33903  
Qy  
3605 CTGAAGCAGTTTAAAAATAATACAGGGAGAGATCTTACTGTGTGGACATCTCATGAGTGA 3664  
Db  
33904 CTGAAGCAGTTTAAAAATAATACAGGGAGAGATCTTACTGTGTGGACATCTCATGAGTGA 33963  
Qy  
3665 AYRGCACTACTCTCTAAAGAGACTTGTGGCTGTGACAACTGTTTACTTAAATRTC 3724  
Db  
33964 ATGGCATACTCACTCTAAAGGAGACTTGTGGCTGTGACAACTGTTTACTTAAATGTC 34023  
Qy  
3725 AGGCTTATTTACTGAAGGCCAGTGTGCRATCTGTGCACTTTGTGCACTCTTAAACCCAG 3784  
Db  
34024 AGGCTTATTTACTTGAAGGCCAGTGTGCGCATCTGTGCACTTTGTGCACTCTTAAACCCAG 34083  
Qy  
3785 YCNCAATTTCTTCCAGCAATGAAGAAAGATATAATATAAATCTGTCAACAATAATTTCTC 3844  
Db  
34084 CCAATTTCTTCCAGCAATGAAGAAAGATATAAATATAAATCTGTCAACAATAATTTCTC 34143  
Qy  
3845 AAACCTATGCCACTCGAGGGACCTTGTAGAGTTCCTTGTGATCTGATCTCYGACCTTCAAC 3904  
Db  
34144 AAACCTATGCCACTCGAGGGACCTTTTAGAGGTTCTTTGACTGATCCCGACC-TCAAC 34202  
Qy  
3905 TTGTATATCTGATGAAGTTCCTTTGTAGAAAGCACTTCCGAAGYGGGTATGAGTGG 3964  
Db  
34203 TTGTATATCTGATGAAGTTCCTTTGTAGAAAGCACTTCCGAAGYGGGTATGAGTGG 34262  
Qy  
3965 GTCAAGTGAATTAATGAAAGTAAATCCCTTCACTCCAGGAACCTAGTGTCTAGCTR 4024

Db	34263	GTCACTGATAATGGAATACCTTTGAAAGTAATCCCTCACTCCAGGAACCTAGTCTCAGCTA	34322
Qy	4025	GCAGAACTAATAGCTTCATYKGGGCACTAGAAATTAGGAGAGAGAAAGGGYAAATATA	4084
Db	34323	GCAGAACTAATAGCTTCATYKGGGCACTAGAAATTAGGAGAGAGAAAGGGYAAATATA	34382
Qy	4085	TATACAGACTCTRTATATGCTYACCTAGTCTCAGTCCATGCGCCATGMRGCAATATGSAAGAA	4144
Db	34383	TATACAGACTCTRTATATGCTYACCTAGTCTCAGTCCATGCGCCATGMRGCAATATGSAAGAA	34442
Qy	4145	AGGGAATTCCTAACTTCYTGAGRGAACAACCTATCAMACATCAGGAAGCCATTAGGABATTA	4204
Db	34443	AGGGAATTCCTAACTTCYTGAGRGAACAACCTATCAMACATCAGGAAGCCATTAGGABATTA	34502
Qy	4205	TTATYTGCGWGTACAGAACTTACAGAGGTGGAGTCTTACACTGCTGGGGTCTATCANAAA	4264
Db	34503	TTATYTGCGWGTACAGAACTTACAGAGGTGGAGTCTTACACTGCTGGGGTCTATCANAAA	34562
Qy	4265	GGAAAGRAAGGGAATATASAAAGGAATGCGCAAGCAKATATTCAAGCMAAAGAGCTGCA	4324
Db	34563	GGAAAGRAAGGGAATATASAAAGGAATGCGCAAGCAKATATTCAAGCMAAAGAGCTGCA	34622
Qy	4325	AGGCAGGACCTCCATTAGAAATGCTTATTAATCTTCCCTTAGTATAGGCTAATCCCTTC	4384
Db	34623	AGGCAGGACCTCCATTAGAAATGCTTATTAATCTTCCCTTAGTATAGGCTAATCCCTTC	34682
Qy	4385	CGGGAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGAAACCTCACAGG-CAG	4443
Db	34683	CGGGAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGAAACCTCACAGGAGCAG	34742
Qy	4444	TTTTCTCCCTCGGGAGGTTAGCCACTGAAAGGAGGAAATACTTTTGGCTGCAACTAT	4503
Db	34743	TTTTCTCCCTCGGGAGGTTAGCCACTGAAAGGAGGAAATACTTTTGGCTGCAACTAT	34802
Qy	4504	CCAATGGAATTACTTTAAACCCCTTCATCAAACTTTTCACTTAGGCATCGATAGCACCCA	4563
Db	34803	CCAATGGAATTACTTTAAACCCCTTCATCAAACTTTTCACTTAGGCATCGATAGCACCCA	34862
Qy	4564	TCARATGGCCAAATCATTATTTACTGGAACAGGCTTTTCAAACTATCAAGCARATAT	4623
Db	34863	TCARATGGCCAAATCATTATTTACTGGAACAGGCTTTTCAAACTATCAAGCARATAT	34922
Qy	4624	CAGGGCTGTGAATGTGCCRARAATAATCCCTGCTTATCGCCAAAGCTCCTTCAGG	4683
Db	34923	CAGGGCTGTGAATGTGCCRARAATAATCCCTGCTTATCGCCAAAGCTCCTTCAGG	34982
Qy	4684	ABAAACAAPACAGGCCATTAACCTGARAARACTGGCAACTGATTTTACCCACAGCCC	4743
Db	34983	ABAAACAAPACAGGCCATTAACCTGARAARACTGGCAACTGATTTTACCCACAGCCC	35042
Qy	4744	AAACCTCAGGATTTTCAGTATCTACTAGTCTGGGTATATCTTTCACGGGTTGGGCGARAG	4803
Db	35043	AAACCTCAGGATTTTCAGTATCTACTAGTCTGGGTATATCTTTCACGGGTTGGGCGARAG	35102
Qy	4804	GCCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGCACCTAGTTTCATGAAATA	4863
Db	35103	GCCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAGGCACCTAGTTTCATGAAATA	35162
Qy	4864	ATTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGTTTCCAGGCC	4923
Db	35163	ATTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGTTTCCAGGCC	35222
Qy	4924	ACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCATTACACTGCGGCTGA	4983
Db	35223	ACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCATTACACTGCGGCTGA	35282
Qy	4984	AGGCCACAGTCTTCAGGAGGTTCGAGAAAATGAATGAAYACTCAAAGGACATCTAAAA	5043
Db	35283	AGGCCACAGTCTTCAGGAGGTTCGAGAAAATGAATGAAYACTCAAAGGACATCTAAAA	35342
Qy	5044	AGCAAAACCCAGAAAACCCACTCACATGGCCCTGYTCTGTGCTTATAGCCTTTAAAAAGA	5103
Db	35343	AGCAAAACCCAGAAAACCCACTCACATGGCCCTGYTCTGTGCTTATAGCCTTTAAAAAGA	35402
Qy	5104	ATCTGCAACTTTTCCCAAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGCCCC	5163
Db	35403	ATCTGCAACTTTTCCCAAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGCCCC	35462
Qy	5164	TTCATAAACCAATGACCTTGCTTGACCCCAAGACAGCCAACTTAGTTGAGACATCACCT	5223
Db	35463	TTCATAAACCAATGACCTTGCTTGACCCCAAGACAGCCAACTTAGTTGAGACATCACCT	35522
Qy	5224	CCTTAGCCAAATATCAACAGTTCTTTAAACAATTAACAAGAACCTATCCCTGAGAAGAGG	5283
Db	35523	CCTTAGCCAAATATCAACAGTTCTTTAAACAATTAACAAGAACCTATCCCTGAGAAGAGG	35582
Qy	5284	GAAAGAACTATTTCACCCWGTGACATGGTATTAGTCAAGTCCCTTCCTCTTAATTCCC	5343
Db	35583	GAAAGAACTATTTCACCCWGTGACATGGTATTAGTCAAGTCCCTTCCTCTTAATTCCC	35642
Qy	5344	CATCCCTAGATACATCCTCGGAGGAGCCCTACCCAGTCAATTTATCTACCCCAACTGCGG	5403
Db	35643	CATCCCTAGATACATCCTCGGAGGAGCCCTACCCAGTCAATTTATCTACCCCAACTGCGG	35702
Qy	5404	TTAAAGTGGCTGGAGTGGAGTCTTTGGATACATCACACTTTGAGTCAAAATCCTGGATATGC	5463
Db	35703	TTAAAGTGGCTGGAGTGGAGTCTTTGGATACATCACACTTTGAGTCAAAATCCTGGATATGC	35762
Qy	5464	CAAAGGAACCTGAAATCCAGGAGACAACGCTAGCTATTCTGTGTGAACCTCTAGAGGATT	5523
Db	35763	CAAAGGAACCTGAAATCCAGGAGACAACGCTAGCTATTCTGTGTGAACCTCTAGAGGATT	35822
Qy	5524	TGGGCTGCTCTTCAAAACAACCAAGGAGGAAAGTAACATAAATCATAAATCCCCCATG	5583
Db	35823	TGGGCTGCTCTTCAAAACAACCAAGGAGGAAAGTAACATAAATCATAAATCCCCCATG	35881
Qy	5584	GSCTCCCTTATCATATATTTTCTCTKTASTGTSTTTTACCTCTCTCTCACTCTCACTGCA	5643
Db	35882	GSCTCCCTTATCATATATTTTCTCTKTASTGTSTTTTACCTCTCTCTCACTCTCACTGCA	35941
Qy	5644	CCCCCTTCAATGCGCTGTATGACCACTAGTCTCCCTTATCCMAGAGTTTCTATGGAGATG	5703
Db	35942	CCCCCTTCAATGCGCTGTATGACCACTAGTCTCCCTTATCCMAGAGTTTCTATGGAGATG	36001
Qy	5704	CAGGCTCCCGAAATATTATGATGCCCATCGTATAGAGTCTTTTSTAAGGGAAACCCCAAC	5763
Db	36002	CAGGCTCCCGAAATATTATGATGCCCATCGTATAGAGTCTTTTSTAAGGGAAACCCCAAC	36061
Qy	5764	TTCACTGCCCCACACCCATATGCGGCACTGCTATCACTCTGCCACTCTTTGCAATGCAT	5823
Db	36062	TTCACTGCCCCACACCCATATGCGGCACTGCTATCACTCTGCCACTCTTTGCAATGCAT	36121
Qy	5824	GCAAAATCTCATTTATGGACAGGAAAAATGATTAATCTCTAGTTGCTCTGAGGACCTTGA	5883
Db	36122	GCAAAATCTCATTTATGGACAGGAAAAATGATTAATCTCTAGTTGCTCTGAGGACCTTGA	36181
Qy	5884	GTCACTGCTGTGTGACCTTACTTCCCAAACTGGTATGCTCTGATGGGGTGGAGTTCAA	5943
Db	36182	GTCACTGCTGTGTGACCTTACTTCCCAAACTGGTATGCTCTGATGGGGTGGAGTTCAA	36241
Qy	5944	GATCAGGCAAGAGAAAAACATGTAAGAAAGTAATCTCCCAACTCAACGSGGTACATGGC	6003
Db	36242	GATCAGGCAAGAGAAAAACATGTAAGAAAGTAATCTCCCAACTCAACGSGGTACATGGC	36301
Qy	6004	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATAAAGCCCTCGTACCCAT	6063
Db	36302	ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATAAAGCCCTCGTACCCAT	36361
Qy	6064	ACTCGCTGGTAAAGCTATTTAATACCAACCTCTCACTGGGCTCCATGAGGTCTCGGCCAA	6123
Db	36362	ACTCGCTGGTAAAGCTATTTAATACCAACCTCTCACTGGGCTCCATGAGGTCTCGGCCAA	36421
Qy	6124	AAACCTTAACTGTGTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATCCCT	6183
Db	36422	AAACCTTAACTGTGTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATCCCT	36481

QY 6184 GTACCTGAACAAATGGAACAACTTTCAGCACAGAAAAATAAACACACACTTCGTTTTAGTAGGA 6243  
Db 36482 GTACCTGAACAAATGGAACAACTTTCAGCACAGAAAAATAAACACACACTTCGTTTTAGTAGGA 36541  
QY 6244 CTTCTGTTTCCAAATGGAATAACCCATACCTCAAACTCCTCAGCTGTGTAAATTTAGC 6303  
Db 36542 CTTCTGTTTCCAAATGGAATAACCCATACCTCAAACTCCTCAGCTGTGTAAATTTAGC 36601  
QY 6304 AATACTACATACACAAACCACTCCCAATGCATCAGGTGGGTAACTCCTCCCAACAAATA 6363  
Db 36602 AATACTACATACACAAACCACTCCCAATGCATCAGGTGGGTAACTCCTCCCAACAAATA 36661  
QY 6364 GTCTGCTACCTCAGGAATAATTTTGTCTGTGTGTACCTCAGCTATGCTGTGTTTGAAT 6423  
Db 36662 GTCTGCTACCTCAGGAATAATTTTGTCTGTGTGTACCTCAGCTATGCTGTGTTTGAAT 36721  
QY 6424 GGCTCTTCAGAAATCTATGTGCTTCCTCTCATTTCTTAGTGCCCCYATGRCCAATCTACACT 6483  
Db 36722 GGCTCTTCAGAAATCTATGTGCTTCCTCTCATTTCTTAGTGCCCCYATGRCCAATCTACACT 36781  
QY 6484 GAACAAGATTTATACAGTTATGTCTATCTAAGCCCGCGAAACAAAGAGTACCCATTCTT 6543  
Db 36782 GAACAAGATTTATACAGTTATGTCTATCTAAGCCCGCGCAACAAAGAGTACCCATTCTT 36841  
QY 6544 CTTTTGTTTATAGGACGAGGTGCTAGTGCACCTAGTGTGCTAGTGGCAATGGCGGTATCACA 6603  
Db 36842 CTTTTGTTTATAGGACGAGGTGCTAGTGCACCTAGTGTGCTAGTGGCAATGGCGGTATCACA 36901  
QY 6604 ACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGGTC 6663  
Db 36902 ACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGGTC 36961  
QY 6664 GCGACTCCTGGTCACTTGAAGATCACTTAACCTCCTAGCAGCAGTAGTCTCTTCA 6723  
Db 36962 GCGACTCCTGGTCACTTGAAGATCACTTAACCTCCTAGCAGCAGTAGTCTCTTCA 37021  
QY 6724 AATCGAAGCTTTAGACTGCTTAACCGCTGAGAGGGGGAACCTGTTTATTTAGG 6783  
Db 37022 AATCGAAGCTTTAGACTGCTTAACCGCTGAGAGGGGGAACCTGTTTATTTAGG 37081  
QY 6784 GAAGATGCTGTTTATTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 6843  
Db 37082 GAAGATGCTGTTTATTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 37141  
QY 6844 CSAGATCGAATAACAGTAKAGCAGARGAGCTTGGAAACACTGGACCTCGGGCCCTCTC 6903  
Db 37142 CGAGATCGAATAACAGTAKAGCAGARGAGCTTGGAAACACTGGACCTCGGGCCCTCTC 37201  
QY 6904 AGCCATGATGCTGCTGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATATATTTGCTA 6963  
Db 37202 AGCCATGATGCTGCTGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATATATTTGCTA 37261  
QY 6964 CTCCTCTTTGGACCTGTTATCTTTACCTCTTTTAACTTTTGTCTCTTCCAGAAATCGAA 7023  
Db 37262 CTCCTCTTTGGACCTGTTATCTTTAACTCTTTTAACTTTTGTCTCTTCCAGAAATCGAA 37321  
QY 7024 GCTGTAAACTACAAATGAGGCCCAAGATCCAGTCCAGACTAAGATCTACCCGAGACC 7083  
Db 37322 GCTGTAAACTACAAATGAGGCCCAAGATCCAGTCCAGACTAAGATCTACCCGAGACC 37381  
QY 7084 CTGGACCGGCTGTTAGCCCAAGATCTGATGTTTAAATGACATCAAGGCGACCTCTCTGAG 7143  
Db 37382 CTGGACCGGCTGTTAGCCCAAGATCTGATGTTTAAATGACATCAAGGCGACCTCTCTGAG 37441  
QY 7144 GAAATCTAGCTGACAACTCTACTACGCCCAATTTCCAGCAGGAAGAGTTAGCGGT 7203  
Db 37442 GAAATCTAGCTGACAACTCTACTACGCCCAATTTCCAGCAGGAAGAGTTAGCGGT 37501  
QY 7204 SGTGCGCCCACTCCCAACAGCACTTAGGTTTCTGTTGATGGGGGAGCTGAGAGAC 7263  
Db 37502 SGTGCGCCCACTCCCAACAGCACTTAGGTTTCTGTTGATGGGGGAGCTGAGAGAC 37561  
QY 7264 AGGACTAGCTGGATTTCTTAGGCTGATTAAGATCCYTAAGCCTAGCTAGGTGACC 7323

Db 37562 AGGACTAGCTGATTTCTTAGGCTAGCTAAGATCCCTAAGCCTAGCTGGGAGGTGACC 37621  
QY 7324 ACATCCACCTTTAAACAGCGGGCTTGCNACTTAGTGTCAACCTGACCAATCAGAGAGCTC 7383  
Db 37622 ACATCCACCTTTAAACAGCGGGCTTGCNACTTAGTGTCAACCTGACCAATCAGAGAGCTC 37681  
QY 7384 ACTAAATGCTAAATTAGGCAAGACAGGAGTAAAGAAATAGCCAAATCATCTATTGCTG 7443  
Db 37682 ACTAAATGCTAAATTAGGCAAGACAGGAGTAAAGAAATAGCCAAATCATCTATTGCTG 37741  
QY 7444 AGAGCACAGCAGGAGGAGCAATGATCGGATATATAACCCAAAGTGTTCAGCGCGCAACGG 7503  
Db 37742 AGAGCACAGCAGGAGGAGCAATGATCGGATATATAACCCAAAGTGTTCAGCGCGCAACGG 37801  
QY 7504 CAACCCCTTTGGGTCCCTCCCTTTGATGGGAGCTCTGTTTTCATGCTATTTCACCTCT 7563  
Db 37802 CAACCCCTTTGGGTCCCTCCCTTTGATGGGAGCTCTGTTTTCATGCTATTTCACCTCT 37861  
QY 7564 ATTAAATCTTCGACCTGCR 7582  
Db 37862 ATTAAATCTTCGACCTGCA 37880

RESULT 5

US-11-028-539-3

; Sequence 3, Application US/11028539

; Publication No. US20050118573A1

; GENERAL INFORMATION:

; APPLICANT: PERIN, Jean-Pierre

; APPLICANT: RIEGER, Francois

; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH H

; FILE OF INVENTION: ENDOGENOUS RETROVIRAL MOTIFS

; FILE REFERENCE: 200936USOPT

; CURRENT APPLICATION NUMBER: US/11/028,539

; CURRENT FILING DATE: 2005-01-05

; PRIOR APPLICATION NUMBER: US/09/719,554

; PRIOR FILING DATE: 2001-01-18

; PRIOR APPLICATION NUMBER: PCT/FR99/01513

; PRIOR FILING DATE: 1999-06-25

; NUMBER OF SEQ ID NOS: 122

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3

; LENGTH: 10499

; TYPE: DNA

; ORGANISM: Homo sapiens

US-11-028-539-3

Query Match 85.9%; Score 6436; DB 13; Length 10499;

Best Local Similarity 95.7%; Pred. No. 0;

Matches 6680; Conservative 195; Mismatches 50; Indels 54; Gaps 28;

QY 606 GATGGGAACGTTCCCGCAAGACAAAACGCCCCCTAAGACGTTATTTGGARAAATTTGGA 665  
Db 2953 GATGGGAACGTTCCCGCAAGACAAAACGCCCCCTAAGACGTTATTTGGARAAATTTGGA 3012  
QY 666 MCAATTTGACCCCTCAGACACTAAGAAAGAAACGACTTATTTCTTCAGTGCCGCTG 725  
Db 3013 CCAATTTGACCCCTCAGACACTAAGAAAGAAACGACTTATTTCTTCAGTGCCGCTG 3072  
QY 726 GCATCTCTGAGGAAGTAAATTAATAACCATCTTACAGCTAGACYTCTTTTGTAGAA 785  
Db 3073 GCATCTCTGAGGAAGTAAATTAATAACCATCTTACAGCTAGACYTCTTTTGTAGAA 3132  
QY 786 AAGGCAATGAGTGAAGTGCATTAAGTACAACTTTCTTTTCAATTAAGAGACAACCTCAC 845  
Db 3133 AAGGCAATGAGTGAAGTGCATTAAGTACAACTTTCTTTTCAATTAAGAGACAACCTCAC 3192  
QY 846 AATATGTAAGGAAGTGAATTTATGCCCTTACAGGAAGCCTTCAGAGTCTACCTCCCTATC 905  
Db 3193 AATATGTAAGGAAGTGAATTTATGCCCTTACAGGAAGCCTTCAGAGTCTACCTCCCTATC 3252





Db 5384 CCTCGCTACCTGTGGCTACATGGTTTCCAAACCAAGGCTCAACTCTGCTCACAGCAGGT 5443  
Qy 3125 TAAATACCTTAGGCTTAARATTAATCAAAAGGCACCAAGGCCCTCAGTGAGGAAYRATCCA 3184  
Db 5444 ----TACTTAGGGCTAARATTAATCAAAAGGCACCAAGGCCCTCAGTGAGGAACATCCA 5499  
Qy 3185 GCCTATACCTGGCTTATCTCATCYCAAAACCCCTAAAGCAACTAAGRRRTTCTCTTGGGRT 3244  
Db 5500 GCCTATACCTGGCTTATCTCATCCCAAAACCCCTAAGCAACTAAGGGGATCTCTTGGCGT 5559  
Qy 3245 AAYAGGYTTCGCCGAANATGATTCGCCAGGTWGTGGRRRAAATAGCCAGGCAATTAATA 3304  
Db 5560 AATAGGTTTCTGCCGAATATGGAAT-CCCAGGTATGGGGAATAGCCAGGTCATTAATA 5618  
Qy 3305 CASTAATTAAGGAACCTCAGAAAGCAATACCCATTTARTAGATGGAYMCTGAAGYMR 3364  
Db 5619 CACTAATTAAGGAACCTCAGAAAGCAATACCCATTTARTAGATGGCAACTGAAGTAG 5678  
Qy 3365 AAGTGGCTTTCAGGCCCTTAAAGGAAGGCCCTTAAACCCCAAGYCCAGGTGTTAAGYTTGCC 3424  
Db 5679 AAGTGGCTTTCAGGCC-----TAACCCCAAGCCCGAGTGTAAAGTTTGGC 5724  
Qy 3425 AACRGGGCAAGACTTTTSTYATAYRTCAAGAAAAAACAAGAAAYAGCTCTRGAGTCC 3484  
Db 5725 AACAGGGCAAGACTTTTCTTCATATGTACAG-AAAAAACAGGAATAGCTTAGGAGTCC 5783  
Qy 3485 TTACACAGTTCRAGGAYGAGCTTGCACACCTGTGGCATACCTGTGGCATACCTGAAGAAATGATG 3544  
Db 5784 TTACACAGATCCGAGGGATGAGCTTGCACACCTGTGGCATACCTGAAGAAATGATG 5843  
Qy 3545 TAGTGGCAAGGTTGRCVYCATGTGTTAYGGGTAGTGGTGCGAGTAGCAGYKTAGTAT 3604  
Db 5844 TAGTGGCAAGGTTGACCTCATGTGTACGGGTAGTGGTGCGAGTAGCAGTCTTAGTAT 5903  
Qy 3605 CTGAAGCAGTTAAATAATACAGGGRAGAGATCTTACTGTGGGACATCTCATGAKGTGA 3664  
Db 5904 CTGAAGCAGTTAAATAATACAGGGAAGAGATCTTACTGTGGGACATCTCATGATGTGA 5963  
Qy 3665 AYRGACTACTGCTAAGAGAGACTTGTGGCTGTGAGCAACAYGTTTACTTAATRTC 3724  
Db 5964 ATGGCATACTCATGCTGAAGGAGACTTGTGGCTGTGAGCAACATGTTTACTTAAATGTC 6023  
Qy 3725 AGGCTCTATTACTTGAAGGGCCAGTGTGCTGACCTGTGCACTTGTGCAACTCTTAAACCCAG 3784  
Db 6024 AGGCTCTATTACTTGAAGGGCCAGTGTGCTGACCTGTGCAACTCTTAAACCCAG 6083  
Qy 3785 YNCATTTCTTCAGACAAATGAAGAAAGATARAAYATACTGTCAACAARTAAATTTCTC 3844  
Db 6084 CCACATTTCTTCAGACAAATGAAGAAAGATARAACATAAATCTGTCAACAAGTAAATTTCTC 6143  
Qy 3845 AAACCTATGCCACTCGAGGGGACCTTGTAGARGTTCCYTTGACTGATCCYGACCTTCAAC 3904  
Db 6144 AAACCTATGCCACTCGAGGGGACCTTGTAGARGTTCCYTTGACTGATCCCGACC-TCAAC 6202  
Qy 3905 TTGTATACTGATGGAGTTCCTTGTAGAAAAGGACTTCGAAAGYGGGGTATCGATG 3964  
Db 6203 TTGTATACTGATGGAGTTCCTTGTAGAAAAGGACTTCGAAAAGTGGGGTATCGATG 6262  
Qy 3965 GTCAGTGAATAATGGGAATAVTTGAAAGTAATCCCTCACTCCAGGAACCTAGTGTAGCTR 4024  
Db 6263 GTCAGTGAATAATGGGAATACTTGAAGTAATCCCTCACTCCAGGAACCTAGTGTAGCTR 6322  
Qy 4025 GCAGAACTAATAGCCYTCATYTKGGGCACTAGAAATTAGGAGAGRAAAAAGGGYAAATATA 4084  
Db 6323 GCAGAACTAATAGCCCTCATCTTGGGCACTAGAAATTAGGAGAGAAAAAGGGCAATATA 6382  
Qy 4085 TATAAGACTCTEARTACTGTACCTAGTCNTCCATGCCCATGMBGCATATGSRAGAA 4144  
Db 6383 TATAAGACTCTAATAATATGCTTACCTAGTCCCTCATGCGCAATATGGAAAGAA 6442  
Qy 4145 AGGGAATTCCTAACTTCYAGRGAACAACCTATCAMAATCAGGAAGCCATTAGGARATTA 4204  
Db 6443 AGGGAATTCCTAACTTCYAGRGAACAACCTATCAMAATCAGGAAGCCATTAGGARATTA 6502

Qy 4205 TTAYTGGCWGTACAGAAACCTARAGAGTGTGMAGTCTTTACATCGYGGGGTCAATCANAAA 4264  
Db 6503 TTATTTGGCTGTACAGAAACCTAAGAGGTGSCAGTCTTACACTGCCGGGGTCAACAGAA 6562  
Qy 4265 GGAAGRAAAGGGGAAATASAGRGAAVTCGCAACAKATATTTGAAGCMAAAGAGCTGCA 4324  
Db 6563 GGAAGRAAAGGGGAAATAGAGAGAACTGCGAAGCAGATATTTGAAGCMAAAGAGCTGCA 6622  
Qy 4325 AGGCAAGGACCTCCATTTAGAAATGCTTTAAACCTTCCCTTTAGTATAGGTTAACTCCCTTC 4384  
Db 6623 AGGCAAGGACCTCCATTTAGAAATGCTTTAAACCAACCCCTTAGTATAGGTTAACTCCCTTC 6682  
Qy 4385 CGGGAACCAAGCCCGAGTACTCAGAGGAGGAAAAAGAAACAGAACTCTCAGAGG-CAG 4443  
Db 6683 CGGGAACCAAGCCCGAGTACTCAGAGGAGGAAAAAGAAACAGAACTCTCAGAGGAGAG 6742  
Qy 4444 TTTTCTCCCTCGGACGGTTCAGCACTGAAGAAAGGAAAAATACTTTTTCCTGCAACTAT 4503  
Db 6743 TTTTCTCCCTCGGACGGTTCAGCACTGAAGAAAGGAAAAATACTTTTTCCTGCAACTAT 6802  
Qy 4504 CCAATGGAATTAATCTTAAACCCCTTCATCAACCTTTTCACTTAGGCATCGATAGCACCA 4563  
Db 6803 CCAATGGAATTAATCTTAAACCCCTTCATCAACCTTTTCACTTAGGCATCGATAGCACCA 6862  
Qy 4564 TCARATGGCCAAATCATTTTACTTGGACAGGCTTTTCAAAAATACTCAAGCARATAT 4623  
Db 6863 TCARATGGCCAAATCATTTTACTTGGACAGGCTTTTCAAAAATACTCAAGCARATAT 6922  
Qy 4624 CAGGCGCTGTGAATGTGCCARABAAATAAATCCCTGCTTATCGCAGGCTCTTTCAGG 4683  
Db 6923 CAGGCGCTGTGAATGTGCCAGAGAAATAAATCCCTGCTTATCGCAGGCTCTTTCAGG 6982  
Qy 4684 ARAACAAARACAGGCCATTTACCTGARAARABACTTGGCAACTGATTTTACCAAGGCC 4743  
Db 6983 AGAACAAAGAACAGGCCATTTACCTGGAAGAGACTGGCAACTGATTTTACCAAGGCC 7042  
Qy 4744 AAACCTCAGGATTTTCACTACTTCTAGTCTGGGTARATATCTTTTACGGGTTGGCAGAG 4803  
Db 7043 AAACCTCAGGATTTTCACTACTTCTAGTCTGGGTARATATCTTTTACGGGTTGGCAGAG 7102  
Qy 4804 GCCTTCCCTCTAGGACAGAAAAAGGCCCAAGAGGTAAATAAAGGCACTAGTTTCAATGAATA 4863  
Db 7103 GCCTTCCCTCTAGGACAGAAAAAGGCCCAAGAGGTAAATAAAGGCACTAGTTTCAATGAATA 7162  
Qy 4864 ATTCCAGATTCGAGACTTCCCGAGGCTTACAGAGTGAACAATAGCCCTGCTTTCAGGCC 4923  
Db 7163 ATTCCAGATTCGAGACTTCCCGAGGCTTACAGAGTGAACAATAGCCCTGCTTTCAGGCC 7222  
Qy 4924 ACAGTTAACCCAGGAGTATCCAGGCGTTAGTATACGATATCACTTACACTGCGCCCTGA 4983  
Db 7223 ACAGTTAACCCAGGAGTATCCAGGCGTTAGTATACGATATCACTTACACTGCGCCCTGA 7282  
Qy 4984 AGGCCACAGTCTCTCAGGGAAGGTCCAGAAAAATGAATAAAGTCAAAAGGACATCTAAAA 5043  
Db 7283 AGGCCACAGTCTCTCAGGGAAGGTCCAGAAAAATGAATAAAGTCAAAAGGACATCTAAAA 7342  
Qy 5044 AAGCAAAACCCAGGAAAAACCCACTCACTGCGCTGTCTGTTGCCCTATAGCCTTAAAAAGA 5103  
Db 7343 AAGCAAAACCCAGGAAAAACCCACTCACTGCGCTGTCTGTTGCCCTATAGCCTTAAAAAGA 7402  
Qy 5104 ATCTGCACTTTTCCCAAAAGGAGGACTTAGCCCATAGCAATGCTCTATGGAAGGCC 5163  
Db 7403 ATCTGCACTTTTCCCAAAAGGAGGACTTAGCCCATAGCAATGCTCTATGGAAGGCC 7462  
Qy 5164 TTCTAAACCAATAGCTTGTGCTTGACCCCAAGACAGCAACTTAGTTGCAAGCATCACCT 5223  
Db 7463 TTCTAAACCAATAGCTTGTGCTTGACCCCAAGACAGCAACTTAGTTGCAAGCATCACCT 7522  
Qy 5224 CTTTAGCCAAATATCAACAGTCTCTTAAAAATTAACAAGGAACCTTATCCCTGAGAGAGG 5283  
Db 7523 CTTTAGCCAAATATCAACAGTCTCTTAAAAATTAACAAGGAACCTTATCCCTGAGAGAGG 7582



QY 5284 GAAAGAACTATTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAATTC 5343  
DB 7583 GAAAGAACTATTCCACCCCTTGTGACATGGTATTAGTCAAGTCCCTTCCTTAATTC 7642  
QY 5344 CATCCCTAGATACATCTCGGAGAGACCCCTACCCAGTCAATTTATYATACCCCACTCGG 5403  
DB 7643 CATCCCTAGATACATCTCGGAGAGACCCCTACCCAGTCAATTTATCTACCCCACTCGG 7702  
QY 5404 TTAAGTGGCTGGAGTGGAGTCTTGGATACATCACTTGGATCAAAATCTCGGATATGC 5463  
DB 7703 TTAAGTGGCTGGAGTGGAGTCTTGGATACATCACTTGGATCAAAATCTCGGATATGC 7762  
QY 5464 CAAAGAACTGAAATCCAGAGAGACAACTGCTAGTATCTCTGAACTCTAGAGATT 5523  
DB 7763 CAAAGAACTGAAATCCAGAGAGACAACTGCTAGTATCTCTGAACTCTAGAGATT 7822  
QY 5524 TGGCCTGCTCTTCAAAACAACAGAGAGAAAGTAACTAAATCAATAATCCCCATG 5583  
DB 7823 TGGCCTGCTCTTCAAAACAACAGAGAGAAAGTAACTAAATCAATAATCCCCATG 7881  
QY 5584 GSCTCCCTTATCATATTTTCTCTTASTGTTSTTTTACCTSTTTTCACTCTCATGCA 5643  
DB 7882 GSCTCCCTTATCATATTTTCTCTTACTGTTTCTTTTACCTCTTTTCACTCTCATGCA 7941  
QY 5644 CCCCTCCATGCCGCTGTATGACCACTAGTCTCCCTTACCAAGAGTCTTCTAGGAGATG 5703  
DB 7942 CCCCTCCATGCCGCTGTATGACCACTAGTCTCCCTTACCAAGAGTCTTCTAGGAGATG 8001  
QY 5704 CAGCGTCCCGAAATATTGATGCCCCATCGTATAGAGTCTTTSTAAGGAAACCCCAAC 5763  
DB 8002 CAGCGTCCCGAAATATTGATGCCCCATCGTATAGAGTCTTTSTAAGGAAACCCCAAC 8061  
QY 5764 TTCACTGCCCAACCCATATGCGCGCAACTGCTATCACTCTGCGCATCTTTGCGATGCAT 8121  
DB 8062 TTCACTGCCCAACCCATATGCGCGCAACTGCTATCACTCTGCGCATCTTTGCGATGCAT 8121  
QY 5824 GCAAACTACTTATTTGGACAGGAAATCAATTAATCTAGTGTCTGAGAGACTTGA 5883  
DB 8122 GCAAACTACTTATTTGGACAGGAAATCAATTAATCTAGTGTCTGAGAGACTTGA 8181  
QY 5884 GTCACTGTCTGTGGACTTACTTCAACCAACTGCTATGCTGATGCGGGGTGGATTCAA 5943  
DB 8182 GTCACTGTCTGTGGACTTACTTCAACCAACTGCTATGCTGATGCGGGGTGGATTCAA 8241  
QY 5944 GATCAGGCAAGAGAAAACAATGTAAGAGAGTAACTCTCCAACTCAACCGGGTACATGGC 6003  
DB 8242 GATCAGGCAAGAGAAAACAATGTAAGAGAGTAACTCTCCAACTCAACCGGGTACATGGC 8301  
QY 6004 ACCTTAGCCCTTACAAAGACTAGATCTCTCAAAACTACATGAAACCCCTCGTACCCAT 6063  
DB 8302 ACCTTAGCCCTTACAAAGACTAGATCTCTCAAAACTACATGAAACCCCTCGTACCCAT 8361  
QY 6064 ACTCGCTGTAGGCTTATTAATACCACTCACTGGGCTCCATGAGGTCTCGGCCAA 6123  
DB 8362 ACTCGCTGTAGGCTTATTAATACCACTCACTGGGCTCCATGAGGTCTCGGCCAA 8421  
QY 6124 AACCTTACTAATGTTGGATATGCTCCCTGAACTTCAAGCCATATGTTTCAATCCCT 6183  
DB 8422 AACCTTACTAATGTTGGATATGCTCCCTGAACTTCAAGCCATATGTTTCAATCCCT 8481  
QY 6184 GTACCTGAAACATGGAACAACTTACGACAGAAATAAACAACCTTCGTTTATGAGGA 6243  
DB 8482 GTACCTGAAACATGGAACAACTTACGACAGAAATAAACAACCTTCGTTTATGAGGA 8541  
QY 6244 CCTCTGTTTCCAACTGGAATAAACCCTACCTCAAACTGCTGTAATAATTTAGC 6303  
DB 8542 CCTCTGTTTCCAACTGGAATAAACCCTACCTCAAACTGCTGTAATAATTTAGC 8601  
QY 6304 AATACTACATACAAACCACTCCCAATGATCAGGTGGGTAACTCTCTCCCAACAAATA 6363  
DB 8602 AATACTACATACAAACCACTCCCAATGATCAGGTGGGTAACTCTCTCCCAACAAATA 8661  
QY 6364 GTCTGCTACCTCAGGAATAATTTTGTGTGTGCTACCTCAGCCTATCGTTGTTGAAT 6423

DB 8662 GTCTGCTACCTCAGGAATAATTTTGTGTGTGCTACCTCAGCCTATCGTTGTTGAAT 8721  
QY 6424 GGCTCTTCAAGAACTATGTGCTTCTCTCAATCTTAGTGCCCCYATGRCATCTACACT 6483  
DB 8722 GGCTCTTCAAGAACTATGTGCTTCTCTCAATCTTAGTGCCCCYATGRCATCTACACT 8781  
QY 6484 GAAACAAGATTTATACAGTTATGTATATCTAAGCCCGCAACAAAGAGTACCATTTCT 6543  
DB 8782 GAAACAAGATTTATACAGTTATGTATATCTAAGCCCGCAACAAAGAGTACCATTTCT 8841  
QY 6544 CCTTTTGTATAGGACGAGGTGCTAGTGCACTAGTACTGCGATTTGCGGTATCACA 6603  
DB 8842 CCTTTTGTATAGGACGAGGTGCTAGTGCACTAGTACTGCGATTTGCGGTATCACA 8901  
QY 6604 ACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGTC 6663  
DB 8902 ACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGTC 8961  
QY 6664 GCGCACTCCCTGCTCACTTGTCAAGATCAACTTAACTCCCTAGCAGCAGTATCTTCTA 6723  
DB 8962 GCGCACTCCCTGCTCACTTGTCAAGATCAACTTAACTCCCTAGCAGCAGTATCTTCTA 9021  
QY 6724 AATCGAAGACTTTAGACTTGTAACTGCTAAGCGCTGAGAGAGGGGAACTGTTTATTTAGG 6783  
DB 9022 AATCGAAGACTTTAGACTTGTAACTGCTAAGCGCTGAGAGAGGGGAACTGTTTATTTAGG 9081  
QY 6784 GAAGATGCTGTTATTTATGTTAATCAATCCGAACTGCTCACTGAGAAAGTTAAAGAAAT 6843  
DB 9082 GAAGATGCTGTTATTTATGTTAATCAATCCGAACTGCTCACTGAGAAAGTTAAAGAAAT 9141  
QY 6844 CSAGATCGAATAACAAGTAKAGCAGAGACTTTCGAAACACTGCACTCCCTGGGCTCTCTC 6903  
DB 9142 CGAGATCGAATAACAAGTAKAGCAGAGACTTTCGAAACACTGCACTCCCTGGGCTCTCTC 9201  
QY 6904 AGCCATGAGTGGCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATAATTTGCTA 6963  
DB 9202 AGCCATGAGTGGCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATAATTTGCTA 9261  
QY 6964 CTCTCTTTGGACCTCTGATCTTTTACCTCTCTGTTAACTTTGCTCTCTCCAGAACTGAA 7023  
DB 9262 CTCTCTTTGGACCTCTGATCTTTTAACTCTCTGTTAACTTTGCTCTCTCCAGAACTGAA 9321  
QY 7024 GCTGTAAACTACTCAAAATGGAGCCCAAGATGCACTCAAGACTAAGATCTACCGCAGACCC 7083  
DB 9322 GCTGTAAACTACTCAAAATGGAGCCCAAGATGCACTCAAGACTAAGATCTACCGCAGACCC 9381  
QY 7084 CTGAGCCGGCTGTTAGCCACGATCTGATGTTAATGAATCAAAAGGCACTCCCTCTGAG 7143  
DB 9382 CTGAGCCGGCTGTTAGCCACGATCTGATGTTAATGAATCAAAAGGCACTCCCTCTGAG 9441  
QY 7144 GAAATCTCAGCTGCACTCACTCTACGCCCAATTCAGCAGGAGCAGTTAGAGGCT 7203  
DB 9442 GAAATCTCAGCTGCACTCACTCTACGCCCAATTCAGCAGGAGCAGTTAGAGGCT 9501  
QY 7204 SGTGGCCCAACCTCCCAACAGCACTTAGGTTTTCTGTTGAGATGGGGACTGAGAGAC 7263  
DB 9502 C-TGGGCCCAACCTCCCAACAGCACTTAGGTTTTCTGTTGAGATGGGGACTGAGAGAC 9560  
QY 7264 AGGACTAGCTGATTTCTTAGGCTGATTAAGAACTCTTAAGCCTTAGTGGGAAGTGA 7323  
DB 9561 AGGACTAGCTGATTTCTTAGGCTGATTAAGAACTCTTAAGCCTTAGTGGGAAGTGA 9620  
QY 7324 ACATCCACCTTTAAACAAGGCTTGCACCTTAGTCTACACCTGACCAATTCAGAGAGCTC 7383  
DB 9621 ACATCCACCTTTAAACAAGGCTTGCACCTTAGTCTACACCTGACCAATTCAGAGAGCTC 9680  
QY 7384 ACTAAAATGCTAATTTAGGCAAGAGAGAGGAGTAAAGAAATAGCCATCATTTATTCMTG 7443  
DB 9681 ACTAAAATGCTAATTTAGGCAAGAGAGAGGAGTAAAGAAATAGCCATCATTTATTCCTG 9740  
QY 7444 AGAGCAGCAGCAGGAGGCAATGATGCGGATATAAACCCCAAGTTCAGAGCCGCAACCG 7503

Db 9741 AGAGCACAGCAGGAGGCAATGATCGGATATAAACCAAGTCTTCGAGCGGCAACGG 9800  
Qy 7504 CAACCCCTTTGGGTCCTCCCTTCTGATGGAGCTCTGTTTCATGCTATTTCACCTCT 7563  
Db 9801 CAACCCCTTTGGGTCCTCCCTTCTGATGGAGCTCTGTTTCATGCTATTTCACCTCT 9860  
Qy 7564 ATTAATCTTGCARCTGCR 7582  
Db 9861 ATTAATCTTGCARCTGCA 9879

RESULT 6  
US-09-854-867-21  
; Sequence 21, Application US/09854867  
; Publication No. US20030224356A1  
; GENERAL INFORMATION:  
; APPLICANT: JOAN, KNOLL H  
; APPLICANT: ROGAN, PETER K  
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING  
; FILE REFERENCE: 30307  
; CURRENT APPLICATION NUMBER: US/09/854,867  
; CURRENT FILING DATE: 2003-05-08  
; NUMBER OF SEQ ID NOS: 613  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 21  
; LENGTH: 8523  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: repeat region  
; LOCATION: (1)-(8523)  
; OTHER INFORMATION: nerv17  
US-09-854-867-21

Query Match 81.8%; Score 6124.6; DB 3; Length 8523;  
Best Local Similarity 95.7%; Pred. No. 0;  
Matches 6363; Conservative 186; Mismatches 50; Indels 53; Gaps 27;

Qy 606 GATGGAAACGTTCCCGCAAGCAAAAAACGCCCCTAAGACGTATTCTGGAGAAATGGGA 665  
Db 1923 GATGGAAACGTTCCCGCAAGCAAAAAACGCCCCTAAGACGTATTCTGGAGAAATGGGA 1982  
Qy 666 MCAATTTGACCTCGACACTAAGAAAGAAAGCACTATTATTTCTTCGAGTCCGCTG 725  
Db 1983 CCAATTTGACCTCGACACTAAGAAAGAAAGCACTATTATTTCTTCGAGTCCGCTG 2042  
Qy 726 GCACCTCTGAGGGAAGTATAAATATAACACCATCTTACAGCTAGACATCTTTGTAGAA 785  
Db 2043 GCACCTCTGAGGGAAGTATAAATATAACACCATCTTACAGCTAGACCTCTTTGTAGAA 2102  
Qy 786 AAGGCAAAATGGAGTGAAGTGCCATAGTAACAATCTTTCTTTTCAATAAGAGACAATCAC 845  
Db 2103 AAGGCAAAATGGAGTGAAGTGCCATAGTAACAATCTTTCTTTTCAATAAGAGACAATCAC 2162  
Qy 846 AATTATGTAAGTGTGTTATTTATGCTACAGGAGCCTTCAGAGTCTACCTCCCTATC 905  
Db 2163 AATTATGTAAGTGTGTTATTTATGCTACAGGAGCCTTCAGAGTCTACCTCCCTATC 2222  
Qy 906 CCAGCATCCCGACTCTTTCCCACTAATAAGAGACCCCTTCAACCCCAATGGTCCAA 965  
Db 2223 CCAGCATCCCGACTCTTTCCCACTAATAAGAGACCCCTTCAACCCCAATGGTCCAA 2282  
Qy 966 AAGGAGATGACAAAAGGGTAAACAGTGAACCAAGAGTGGCAATATCCCAATATGGA 1025  
Db 2283 AAGGAGATGACAAAAGGGTAAACAGTGAACCAAGAGTGGCAATATTCCTCCCAATATGGA 2342  
Qy 1026 CCCCTCCAGAGTGGGAGGAGAAATTCGGCCAGCCAGTGCATGTCGCTTTT 1085  
Db 2343 CCCCT-CCAAGCAGTGGGAGGAGAAATTCGGCCAGCCAGGAGTGCATGTCGCTTTTTC 2401  
Qy 1086 TCTCCAGACTTAAAGCAATAAAAAACAGACTTAGGTAAATCTCAGATAAATCTGATGG 1145  
Db 2402 TCTCCAGACTTAAAGCAATAAAAAACAGACTTAGGTAAATCTCAGATAAATCTGATGG 2461

Qy 1146 CTATATTGRTGCTTTTCAAGGGTTAGGACAATCTTTTGATCTGACATCGAGAGATATATA 1205  
Db 2462 CTATATTGATGTTTTACAGGGTTAGGACAATCTTTTGATCTGACATCGAGAGATATA-A 2520  
Qy 1206 TGTCACTGCTAAATCAGACACTAAACCCCAATGAGAGAAGTGCCACCAATACTGCAGCT 1265  
Db 2521 TGTCACTGCTAAATCAGACACTAAACCCCAATGAGAGAAGTGCCACCAATACTGCAGCT 2580  
Qy 1266 GAGRGTTGGCGATCTCTGGTATCTCAGTCAAGTCAATGAGATANGGATGACAAACAAGG 1325  
Db 2581 GAGAGTTTGGCGATCTCTGGTATCTCAGTCAAGTCAATG--ATAGGATGACAACAG-AGG 2637  
Qy 1326 AAAGNAATGATTTCCCAACAGCCGACGARGCAGTTCCCAAGTCTASACCTCATTTGGGAC 1385  
Db 2638 AAAGGAATGATTTCCCAACAGCCGACGARGCAGTTCCCAAGTCTAGACCTCATTTGGG--- 2694  
Qy 1386 ACAGAAATCAGTAACATCGGAGATTTGGTCTGACAGACATTTGCTAACTTTGTGTASAA 1445  
Db 2695 ACACAGATCAGACAT--GGAGATTTGGTCTGACAGACATTTGCTAACTTTGTGTAGNA 2753  
Qy 1446 GGACTAAGGAAAACTASGAGAAAAARTCTAYGAAATTTACTCAATGATGTCCACATACACA 1505  
Db 2754 GGACTAAGGAAAACTAGGAAG-AAAGTCTATGAAATTTACTCAATGATGTCCACCATACACA 2812  
Qy 1506 GGGGAAGGAAGAAATCTTACTGCTTCTGAGAGACTAAGGGAGGCATTTGAGGAAGC 1565  
Db 2813 -GGGAAGGAAGAAATCTTACTGCTTCTGAGAGACTAAGGGAGGCATTTGAGGAAGC 2871  
Qy 1566 GTGCTCTCTGCTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAAGTTTATCA 1625  
Db 2872 GTGCTCTCTGCTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAAAGTTTATCA 2931  
Qy 1626 CTCAGTCAAGTGCAGACATTAG-AAAAAATCTTCAAAAGTCTGCCGTAGGCCCGGAGCAAA 1684  
Db 2932 CTCAGTCAAGTGCAGACATTAGAAAAAATCTTCAAAAGTCTGCCGTAGGCCCGGAGCAAA 2991  
Qy 1685 ACTTAGAAGCCTATTGAACTTGGCAACCTCGGTTTTTATAATAGAGATCAGGAGGAGC 1744  
Db 2992 ACTTAGAAGCCTATTGAACTTGGCAACCTCGGTTTTTATAATAGAGATCAGGAGGAGC 3051  
Qy 1745 AGGCGGAAACAGACAAAACGGGATTAATAAAGGCAACCGCTTTTAGTCAATGACCCCTCAGG 1804  
Db 3052 AGGCGGAAACAGACAAAACGGGATTAATAAAGGCAACCGCTTTTAGTCAATGACCCCTCAGG 3111  
Qy 1805 CAAGTGGACTTTGGAGGCTCTGGAAAAAGGAAAAAGCTGGGCAAAATGAAATGCTTAATAGG 1864  
Db 3112 CAAGTGGACTTTGGAGGCTCTGGAAAAAGGAAAAAGCTGGGCAAAATGAAATGCTTAATAGG 3171  
Qy 1865 GCTTGCTTCCAGTGGGCTACAGGACACTTTTAAAGAGATTTGTCAGTAGAAGTAG 1924  
Db 3172 GCTTGCTTCCAGTGGGCTACAGGACACTTTTAAAGAGATTTGTCAGTAGAAGTAG 3231  
Qy 1925 CGCGCCCTTCCATGCGCTTATTTCAAGGGAATCACTGGAAGGCCACCTGCCCCCAGG 1984  
Db 3232 CGCGCCCTTCCATGCGCTTATTTCAAGGGAATCACTGGAAGGCCACCTGCCCCCAGG 3291  
Qy 1985 GGAACAAAGGTCTTTTGAAGTCAAGGCACTAAACAGATGATCCAGAGAGGAGCTGAGGG 2044  
Db 3292 GGAACAAAGGTCTCTCTGAGTCAAGGCACTAAACAGATGATCCAGAGAGGAGCTGAGGG 3351  
Qy 2045 TGCTTGGGGCAGGCGCATCCCATGCCATCACCTCAGAGGCCCTGGGTATGCTTGACC 2104  
Db 3352 TGCTTGGGGCAGGCGCATCCCATGCCATCACCTCAGAGGCCCTGGGTATGCTTGACC 3411  
Qy 2105 ATTGAGGCGCAGGAAAGTTGCTCTCGGACACTGTGTGGGCTTTCTTAGTCTTACTCTTC 2164  
Db 3412 ATTGAGGCGCAGG-AGGTTGCTCTCTGAGCACTGTGTGGGCTTTCTTAGTCTTACTCTTC 3470  
Qy 2165 TGTCCCAGACAACTGTCTCTCCAGATCTGTACTATACTTGAGGGGGGTCCNTAAGACGGGCA 2224  
Db 3471 TGTCCCAGACAACTGTCTCTCCAGATCTGTACTATACTTGAGGGGGGTCC-ATAAGACGGGCA 3528



Db 5653 CGGGAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGAAACCTCAGCAGGACAG 5712  
Qy 4444 TTTTCTCCCTCGGACGGTATAGCCACTGAAGAAGGAAAATACTTTTGCGCTGCAACTAT 4503  
Db 5713 TTTTCTCCCTCGGACGGCTAGCCACTGAGAGGAGAAAATACTTTTGCGCTGCAACTAT 5772  
Qy 4504 CCAATGGAAATTACTTAAACCCCTTCATCAAAACCTTTCACTTAGGCATCGATGACCCCA 4563  
Db 5773 CCAATGGAAATTACTTAAACCCCTTCATCAAAACCTTTCACTTAGGCATCGATGACCCCA 5832  
Qy 4564 TCARATGGCAATATATTTATTTATCTGGACCGAGCCCTTTTCAAAAATATCAAGCABATAT 4623  
Db 5833 TCAGATGGCAATATATTTATTTACTGACACAGGCCCTTTTCAAAAATATCAAGCABATAT 5892  
Qy 4624 CAGGCGCTGTGAAGTGTGCCABARAAATATCCCTGCGCTATGCCAAGCTCCTTCAGG 4683  
Db 5893 CAGGCGCTGTGAAGTGTGCCABARAAATATCCCTGCGCTATGCCAAGCTCCTTCAGG 5952  
Qy 4684 ARAACAAARAAACAGGCCATTACCTGRRARAARACTGGCAACTGATTTTACCCCAAGGCC 4743  
Db 5953 AGAACAAAGAACAGGCCATTACCTGGAGAAGACTGGCAACTGATTTTACCCCAAGGCC 6012  
Qy 4744 AAACCTCAGGATTTTCAAGTATCTACTAGTCTGGGTARATATCTTCAAGGTTGGGCAAG 4803  
Db 6013 AAACCTCAGGATTTTCAAGTATCTACTAGTCTGGGTARATATCTTCAAGGTTGGGCAAG 6072  
Qy 4804 GCCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCATAGTTCATGAATA 4863  
Db 6073 GCCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCATAGTTCATGAATA 6132  
Qy 4864 ATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTCTTTCAGGCGC 4923  
Db 6133 ATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTCTTTCAGGCGC 6192  
Qy 4924 ACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGA 4983  
Db 6193 ACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCCTGA 6252  
Qy 4984 AGGCCACAGTCTCAGGGAAGTTCGAGAAAATGAAATGAAGAACTCAAGGACATCTAAA 5043  
Db 6253 AGGCCACAGTCTCAGGGAAGTTCGAGAAAATGAAATGAAGAACTCAAGGACATCTAAA 6312  
Qy 5044 AAGCAAAACCCAGAAACCCACTCAGTGCCTGTCTGTGCTATAGCCTTAAAGAA 5103  
Db 6313 AAGCAAAACCCAGAAACCCACTCAGTGCCTGTCTGTGCTATAGCCTTAAAGAA 6372  
Qy 5104 ATCTGCACTTTCCCAAAAGCAGGACTTAGCCATACGAAATGCTGTATGGAAGGCC 5163  
Db 6373 ATCTGCACTTTCCCAAAAGCAGGACTTAGCCATACGAAATGCTGTATGGAAGGCC 6432  
Qy 5164 TTTCATACCAATGACCTTGTGTGACCCAGACGCACTTAGTTCAGACATCACT 5223  
Db 6433 TTTCATACCAATGACCTTGTGTGACCCAGACGCACTTAGTTCAGACATCACT 6492  
Qy 5224 CTTTAGCCAAATATCAACAAAGTTCTTAAACATTAACAGAACTTATCCCTGAGAGAGG 5283  
Db 6493 CTTTAGCCAAATATCAACAAAGTTCTTAAACATTAACAGAACTTATCCCTGAGAGAGG 6552  
Qy 5284 GAAAGAACTATTCCACCCWGTGTGACATGATTAAGTCAAGTCCCTTCCTCTAAATCCC 5343  
Db 6553 GAAAGAACTATTCCACCCWGTGTGACATGATTAAGTCAAGTCCCTTCCTCTAAATCCC 6612  
Qy 5344 CATCCCTAGATATCATCTCGGAGAGCCCTACCCAGTCAATTTATTTACCCCACTGGGG 5403  
Db 6613 CATCCCTAGATATCATCTCGGAGAGCCCTACCCAGTCAATTTATTTACCCCACTGGGG 6672  
Qy 5404 TTAAAGTGGCTGAGTGGAGTCTTGATATCATCATCACTTTGAGTCAAACTCTGATCTGC 5463  
Db 6673 TTAAAGTGGCTGAGTGGAGTCTTGATATCATCATCACTTTGAGTCAAACTCTGATCTGC 6732  
Qy 5464 CAAAGGAACCTGAAATCCAGGAGACACGCTAGCTATTTCTGTGAACTCTAGAGGATT 5523  
Db 6733 CAAAGGAACCTGAAATCCAGGAGACACGCTAGCTATTTCTGTGAACTCTAGAGGATT 6792

Qy 5524 TGGCGCTCTCTTCAAAACAAACAGGAGGAAAGTAACTAAAAATCAATAATCCCCCATG 5583  
Db 6793 TGGCGCTCTCTTCAAAACAAACAGGAGGAAAGTAACTAAAAATCAATAATCCCCCATG 6851  
Qy 5584 GSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTTAAAGGGAAACCCCCACC 5643  
Db 6852 GSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTTAAAGGGAAACCCCCACC 6911  
Qy 5644 CCCCCTCATGCGCTGTATGACACGATAGCTCCCTTACCMAGAGTTTCTATGGAGATG 5703  
Db 6912 CCCCCTCATGCGCTGTATGACACGATAGCTCCCTTACCMAGAGTTTCTATGGAGATG 6971  
Qy 5704 CAGCGTCCCGGAAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGGAAACCCCCACC 5763  
Db 6972 CAGCGTCCCGGAAAATATTGATGCCCATCGTATAGGAGTCTTTTAAAGGGAAACCCCCACC 7031  
Qy 5764 TTCACTGCCACACCCCATATGATGCCGAACTGCTATCACTCTGCACTCTTTGCAATGAT 5823  
Db 7032 TTCACTGCCACACCCCATATGATGCCGAACTGCTATCACTCTGCACTCTTTGCAATGAT 7091  
Qy 5824 GCNAATATCTATTATTTGGACAGGAAAATGATTAATCTAGTGTCTCTGGAGGACTTGA 5883  
Db 7092 GCNAATATCTATTATTTGGACAGGAAAATGATTAATCTAGTGTCTCTGGAGGACTTGA 7151  
Qy 5884 GTCACTGTCTGTGGACTTACTTCAACCCAAACTGGTATGTCTGATGGGGTGGAGTTCAA 5943  
Db 7152 GTCACTGTCTGTGGACTTACTTCAACCCAAACTGGTATGTCTGATGGGGTGGAGTTCAA 7211  
Qy 5944 GATCAGGCAAGAGAAAAACATGTAAAAGAGTAATCTCCCAAACCTCACCSGGGTACATGCG 6003  
Db 7212 GATCAGGCAAGAGAAAAACATGTAAAAGAGTAATCTCCCAAACCTCACCSGGGTACATGCG 7271  
Qy 6004 ACCTTAGCCCTTACAAAGGACTAGATCTCTCAAAAACCTACATGAAAAACCTCGGTACCCAT 6063  
Db 7272 ACCTTAGCCCTTACAAAGGACTAGATCTCTCAAAAACCTACATGAAAAACCTCGGTACCCAT 7331  
Qy 6064 ACTGCGCTGTGAAGCTATTAAATACCAACCTCAGTGGGCTCCATGAGGTCTCGGCCCAA 6123  
Db 7332 ACTGCGCTGTGAAGCTATTAAATACCAACCTCAGTGGGCTCCATGAGGTCTCGGCCCAA 7391  
Qy 6124 AACCTACTAACTGTGGATATGCTCCCTGCTGAACTTCARGGCCATATGTTTCAATCCCT 6183  
Db 7392 AACCTACTAACTGTGGATATGCTCCCTGCTGAACTTCARGGCCATATGTTTCAATCCCT 7451  
Qy 6184 GTA CCGTGAACATGGAACAACTTCAGACAGAAATAAACAACA CACTTCGTTTTAGTAGGA 6243  
Db 7452 GTACCTGAACATGGAACAACTTCAGACAGAAATAAACAACA CACTTCGTTTTAGTAGGA 7511  
Qy 6244 CCTCTGTGTTTCCAATSTGGAATAACCCATACCTCAAAACCTCACCTGTGTAATAATTTAGC 6303  
Db 7512 CCTCTGTGTTTCCAATSTGGAATAACCCATACCTCAAAACCTCACCTGTGTAATAATTTAGC 7571  
Qy 6304 AATACTACATACAAACCAACTCCCAATGCACTCAGTGGGTAACTCTCCCAACACAAATA 6363  
Db 7572 AATACTACATACAAACCAACTCCCAATGCACTCAGTGGGTAACTCTCCCAACACAAATA 7631  
Qy 6364 GTCTGCTTACCCCTCAGGAATAATTTTTGTCTGTGTGTA CTTAGCCCTATGCTTTTGAAT 6423  
Db 7632 GTCTGCTTACCCCTCAGGAATAATTTTTGTCTGTGTGTA CTTAGCCCTATGCTTTTGAAT 7691  
Qy 6424 GGCTCTTCAGGAATCTATGCTTCTCTCAATTTAGTGGCCCTCATGCACTCTACACT 6483  
Db 7692 GGCTCTTCAGGAATCTATGCTTCTCTCAATTTAGTGGCCCTCATGCACTCTACACT 7751  
Qy 6484 GAAACAGATTTATACAGTTATGTCTATATCTAAGCCCGCAACAAAGAGTACCCATTTCT 6543  
Db 7752 GAAACAGATTTATACAGTTATGTCTATATCTAAGCCCGCAACAAAGAGTACCCATTTCT 7811  
Qy 6544 CCTTTTGTATAGGACGAGGAGTGTAGGTGCACTAGGTGCACTAGGTGCACTGGCAATTTGGCGGTATCACA 6603  
Db 7812 CCTTTTGTATAGGACGAGGAGTGTAGGTGCACTAGGTGCACTGGCAATTTGGCGGTATCACA 7871

```

QY 6604 ACCTCTACTGTTCTACTCAAACTATCTCAAGAACTAAATGGGGACATGGACGGGTC 6663
DB 7872 ACCTCTACTGTTCTACTCAAACTATCTCAAGAACTAAATGGGGACATGGACGGGTC 7931
QY 6664 GCCGACTCCCTGGTCACTTGCAGAGTCAAACTTAACTCCCTAGCAGTAGTCTCTTCA 6723
DB 7932 GCCGACTCCCTGGTCACTTGCAGAGTCAAACTTAACTCCCTAGCAGTAGTCTCTTCA 7991
QY 6724 AATCGAAGAGCTTTAGAGCTGCTAAACCGCTGAGAGAGGGGAACTGTTTATTTTAGGG 6783
DB 7992 AATCGAAGAGCTTTAGAGCTGCTAAACCGCTGAGAGAGGGGAACTGTTTATTTTAGGG 8051
QY 6784 GAAGAATGCTGTTTATTTATGTTTAACTCAATCCGGAACTGCTCACTGAGAAAGTTTAAAGAAAT 6843
DB 8052 GAAGAATGCTGTTTATTTATGTTTAACTCAATCCGGAACTGCTCACTGAGAAAGTTTAAAGAAAT 8111
QY 6844 CSAGATCGAATACAACTAGKAGCAGARGAGCTTGCAGAACTGAGACCTCGGGGCTCTCTC 6903
DB 8112 CGAGATCGAATACAACTAGKAGCAGARGAGCTTGCAGAACTGAGACCTCGGGGCTCTCTC 8171
QY 6904 AGCRAATGATCCCTGATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6963
DB 8172 AGCRAATGATCCCTGATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 8231
QY 6964 CTCTCTTTTGGACCTCTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 7023
DB 8232 CTCTCTTTTGGACCTCTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 8291
QY 7024 GCTGTAAACTCAAAATGGAGCCCAAGATGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 7083
DB 8292 GCTGTAAACTCAAAATGGAGCCCAAGATGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 8351
QY 7084 CTGACCGGCTGTTAGCCCAAGTCTGATGTTTAAATGACATCAAAAGGACCCCTCTCTGAG 7143
DB 8352 CTGACCGGCTGTTAGCCCAAGTCTGATGTTTAAATGACATCAAAAGGACCCCTCTCTGAG 8411
QY 7144 GAAATCTCAGTCAAGCACTCTTACTAGCCCAAGTCTGATGTTTAAATGACATCAAAAGGACCCCTCTCTGAG 7203
DB 8412 GAAATCTCAGTCAAGCACTCTTACTAGCCCAAGTCTGATGTTTAAATGACATCAAAAGGACCCCTCTCTGAG 8471
QY 7204 SGTCGGCAACTCCCAAGCACTTACTAGTCTTCTGTTTCTGTTGAGATGGGGAC 7255
DB 8472 CGTCGGCAACTCCCAAGCACTTACTAGTCTTCTGTTTCTGTTGAGATGGGGAC 8523

```

```

RESULT 7
US-10-786-970A-21
; Sequence 21, Application US/10786970A
; Publication No. US200506449A1
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL
; APPLICANT: ROGAN, PETER
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/10/786, 970A
; CURRENT FILING DATE: 2004-02-24
; PRIOR APPLICATION NUMBER: US/09/573,080
; PRIOR FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 8523
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: (1)..(8523)
; OTHER INFORMATION: herv17
; PUBLICATION INFORMATION:
; AUTHORS: Jurka, J; Walichiewicz, J; Milosavljevic, A
; TITLE: Prototypic sequences for human repetitive DNA
; JOURNAL: Journal of Molecular Evolution

```

```

; VOLUME: 35
; ISSUE: 4
; PAGES: 286-291
; DATE: 1992-10-
; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
; DATABASE ENTRY DATE:
; DATABASE ENTRY DATE: 1996-01-26
; US-10-786-970A-21

Query Match      81.8%; Score 6124.6; DB 10; Length 8523;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 6363; Conservative 186; Mismatches 50; Indels 53; Gaps 27;

QY 606 GATGGGAAAGCTTCCCGGCAAGACAAAGACGCCCCCTAAGACGTAATCTGGAATTTGGGA 665
DB 1923 GATGGGAAAGCTTCCCGGCAAGACAAAGACGCCCCCTAAGACGTAATCTGGAATTTGGGA 1982
QY 666 MCAATTTGACCTTCAGACACTTAAGAAAGAAAGACGTTATATCTTCTGAGTGCCTGCTG 725
DB 1983 CCAATTTGACCTTCAGACACTTAAGAAAGAAAGACGTTATATCTTCTGAGTGCCTGCTG 2042
QY 726 GCACTCTCTGAGGAAAGTATAAATTTATAACACCATCTTACAGCTAGACGTCCTTTGTAGAA 785
DB 2043 GCACTCTCTGAGGAAAGTATAAATTTATAACACCATCTTACAGCTAGACGTCCTTTGTAGAA 2102
QY 786 AAGGCAAAATGGAGTGAAGTGCATTAAGTACAACTTCTTTTCAATTAAGAGACAACTCAC 845
DB 2103 AAGGCAAAATGGAGTGAAGTGCATTAAGTACAACTTCTTTTCAATTAAGAGACAACTCAC 2162
QY 846 AATATGTAAGAAAGTGTGATTTATGCTTACAGGAGCCCTTACAGTCTACCTCCCTATC 905
DB 2163 AATATGTAAGAAAGTGTGATTTATGCTTACAGGAGCCCTTACAGTCTACCTCCCTATC 2222
QY 906 CAGACATCCCGACTCTTCCCAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 965
DB 2223 CAGACATCCCGACTCTTCCCAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2282
QY 966 AAGGAGATAGACAAAGGTTAAACAGTGAACCAAGAGTGCCTAATTTCCCAATTTATGA 1025
DB 2283 AAGGAGATAGACAAAGGTTAAACAGTGAACCAAGAGTGCCTAATTTCCCAATTTATGA 2342
QY 1026 CCGCTCCCAAGCAGTGGGAGGAGAGAAATTCGCCCCAGCAGAGTGCATGTGCTTTTYY 1085
DB 2343 CCGCTCCCAAGCAGTGGGAGGAGAGAAATTCGCCCCAGCAGAGTGCATGTGCTTTTTC 2401
QY 1086 TCTCCAGACTTAAAGCAAAATAAACAGACTTAGTAAATTTCTCAGATAAATCTCAGATGCTGCTTTTC 1145
DB 2402 TCTCCAGACTTAAAGCAAAATAAACAGACTTAGTAAATTTCTCAGATAAATCTCAGATGCTGCTTTTC 2461
QY 1146 CTATATTTGTTTGTATTAAGGGTTAGGACAAATTTCTTGTATCTGACATGGAGAGATATA 1205
DB 2462 CTATATTTGTTTGTATTAAGGGTTAGGACAAATTTCTTGTATCTGACATGGAGAGATATA-A 2520
QY 1206 TGTCTCTGCTAAATCAGACACTAAACCCCAATTAAGAGAGAGTGCACCAATTAATTCAGCCT 1265
DB 2521 TGTCTCTGCTAAATCAGACACTAAACCCCAATTAAGAGAGAGTGCACCAATTAATTCAGCCT 2580
QY 1266 GAGGTTTGGCGATCTCTGTTATCTCAGTCAAGTCAATGATGATGATGATGATGATGATGATGATGATG 1325
DB 2581 GAGGTTTGGCGATCTCTGTTATCTCAGTCAAGTCAATGATGATGATGATGATGATGATGATGATGATG 2637
QY 1326 AAAGANAATGATTTCCCAAGCAGCAGCAGTTCCTCCAGTCTTASACCTCTTASACCTCTTASACCTCTTAS 1385
DB 2638 AAAGANAATGATTTCCCAAGCAGCAGCAGTTCCTCCAGTCTTASACCTCTTASACCTCTTASACCTCTTAS 2694
QY 1386 ACAGAAATCAGTAACATGAGAGATTTGTTCTGAGACATTTGCTAACTTTGTTGCTTASAA 1445
DB 2695 ACAGAAATCAGTAACATGAGAGATTTGTTCTGAGACATTTGCTAACTTTGTTGCTTASAA 2753
QY 1446 GCACTAAGGAAATCTAGAGAGAAATCTTAYGAATTTACTCAATGATGATGATGATGATGATGATGATGATG 1505
DB 2754 GCACTAAGGAAATCTAGAGAGAAATCTTAYGAATTTACTCAATGATGATGATGATGATGATGATGATGATG 2812

```



1506 GGGGAAGGGAAGAAATCTACTGCTTTCTGAGAGACTAAGGAGGCGATTGAGGAAGC 1565  
 2813 -GGGAAGGGAAGAAATCTACTGCTTTCTGAGAGACTAAGGAGGCGATTGAGGAAGC 2871  
 1566 GTGCTCTCTGTGACCTGCTCTGAGGCGCAACTAATCTTAAAGCGTAAGTTTATCA 1625  
 2872 GTGCTCTCTGTGACCTGCTCTGAGGCGCAACTAATCTTAAAGCGTAAGTTTATCA 2931  
 1626 CTCAGTCAGCTGACAGACATTAG-AAAAAATCTCAAAAGTCTGCGTAGGCGCGAGCAAA 1684  
 2932 CTCAGTCAGCTGACAGACATTAGAAAAAATCTCAAAAGTCTGCGTAGGCGCGAGCAAA 2991  
 1685 ACTTAGAAACCCCTATTGAACTTGGCAACVTGCGTTTTTATATATAGAGATCAGAGGAGC 1744  
 2992 ACTTAGAAACCCCTATTGAACTTGGCAACCTCGTTTTTATATATAGAGATCAGAGGAGC 3051  
 1745 AGGCGGAAACAGGACAAACGGGATTAAAAAAAAGGCGCACCGCTTTAGTCATGACCCCTCAGG 1804  
 3052 AGGCGGAAACAGGACAAACGGGATTAAAAAAAAGGCGCACCGCTTTAGTCATGACCCCTCAGG 3111  
 1805 CAACTGGACTTTGGAGGCTCTCGAAAGAGGAAAGCTGGGCAAAATGAAATGCTTAATAGG 1864  
 3112 CAACTGGACTTTGGAGGCTCTCGAAAGAGGAAAGCTGGGCAAAATGAAATGCTTAATAGG 3171  
 1865 GCTTGTCTCCAGTGGCGGCTACAGGACACACTTTAAAAAAGATTGTCGAAGTAGAAGTAAG 1924  
 3172 GCTTGTCTCCAGTGGCGGCTACAGGACACACTTTAAAAAAGATTGTCGAAGTAGAAGTAAG 3231  
 1925 CGCGCCCTCTGCTCCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGG 1984  
 3232 CGCGCCCTCTGCTCCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCCAGG 3291  
 1985 GGACAAAGGCTCTTTGAGTCAGAGCACAATAACAGATGATCCAGCAGCAGGACTGAGGG 2044  
 3292 GGACAAAGGCTCTCTGAGTCAGAGCACAATAACAGATGATCCAGCAGCAGGACTGAGGG 3351  
 2045 TGCTGGGGCAAGCGGCATCCCATGCCATCCCTCACAGAGCCCTGGGTATGCTTGACC 2104  
 3352 TGCTGGGGCAAGCGGCATCCCATGCCATCCCTCACAGAGCCCTGGGTATGCTTGACC 3411  
 2105 ATTGAGGGCGAGGAAGTTGTCTCTGGACACTGTGTGCGGTCTTCTTAGTCTTACTCTTC 2164  
 3412 ATTGAGGGCGAGG-AGTTGTCTCTGGACACTGTGTGCGGTCTTCTTAGTCTTACTCTTC 3470  
 2165 TGTCCCGGACAACTGTCCTCAGATCTGTCACTAATCTGAGGGGGTCCNTAAGACGGCA 2224  
 3471 TGTCCCGGACAACTGTCCTCAGATCTGTCACTA-TCTGAGGGGGTCC-TAAGACGGGCA 3528  
 2225 GTCACCTAGATACCTTTTCCAGCCACTAAGTTATGAAGTGGGGGCTTTATTTCTTTCAC 2284  
 3529 GTCACCTAGATAC-TTCTCCAGCCACTAAGTTATG-ACTGGGGGAGCTTTATTTCTTTCAC 3586  
 2285 ATGCTTTTCTAATTATGCTTGAAGGCCCACTACCTCTGTTAGGAGAGACATTTAGCAAA 2344  
 3587 ATGCTTTTCTAATTATGCTTGAAGGCCCACTACCTCTGTTAGGAGAGACATTTAGCAAA 3646  
 2345 AAGCAGGGGCCATTATACACCTGAACATAGGAGAGAAACACCCGTTGTTGTCCTG 2404  
 3647 AAGCAGGGGCCATTATACACCTGAACATAGGAGAGAAACACCCGTTTGTGT-CCCTGT 3705  
 2405 CTTGAGGAGGAATTAATCCTGAGTCTGGGCAACAGAGGACAATATGAGCGGCCAA 2464  
 3706 CTTGAGGAGGAATTAATCCTGAGTCTGGGCAACAGAGGACAATATGAGCGAG-CAAA 3764  
 2465 GAATGCCGCTCTGTTTCAAGTTAACTAAAGGATTCACCTTCTTCCCTACCAAGGCA 2524  
 3765 GAATGCCGCTCTGTTTCAAGTTAACTAAAGGATTCACCTTCTTCCCTACCAAGGCA 3824  
 2525 GTACCCCTCTAGACCCCAAGGCCCAACAGGATTTCCAAAAGATTGTTAAGGACTTAAAGC 2584  
 3825 GTACCCCTCTAGACCCCAAGGCCCAACAGGACTTCCAAAAGATTGTTAAGGACTTAAAGC 3884  
 2585 CCAAGGCTTAGTAAABACCATGCACTACTCCCTGCAGTAATTCCTGAGTGGATTGAGGAGG 2644

3885 CCAAGGCTAGTAAACCATGCACTAACCCCTGAGTACTCC-----AATTTTAGGAG 3937  
 2645 CACAGAAACCCAGTGGGACAGTGGAGGTTAGTCAAGATCTCAGGATTTATCAATGAGGC 2704  
 3938 TACAGAAACCCAAACAGACAGTGGGA-GGTTAGTCAAGATCTCAGGATTTATCAAT-GAGGC 3995  
 2705 GGTGTGCTTTTATACCCAGCTGTACCTTAGCCCTTATATCTGTMVTTTCCCAAAATACCAGA 2764  
 3996 TGTGTTCTCTATAGCCAGCTGACCTAGGCCCTTATCTCTGTTTCCCAATACCAGA 4055  
 2765 GGAACGACAGTGGTTTACASTCCTGGACCTTTMAGGATGCCCTTCTTCTGCATCCCTGTACA 2824  
 4056 GGAACGACAGTGGTTTACAGTCTCTGGACCTTTCAGGATGCCCTTCTTCTGCATCCCTGTACA 4115  
 2825 TCCTGACTCTCAATCTCTGTTTGGCTTTTGAAGATACTTTCAAAACCCARCATCTCAACTAC 2884  
 4116 TCCTGACTCTCAATCTCTGTTTGGCTTTTGAAGATACTTTCAAAACCCARCATCTCAACTAC 4175  
 2885 CTGGACTTTTACCCCAAGGTTTCAGGATAGTCCCATCTATTTGGCCAGGCAATTAGC 2944  
 4176 CTGGACTTTTACCCCAAGGTTTCAGGATAGTCCCATCTATTTGGCCAGGCAATTAGC 4235  
 2945 CCAAGACTTGAGYCARVTMTCATACCTGGACACTTCTTCTCTCRGTAKGTGGATGATTT 3004  
 4236 CCAAGACTTGAGCAACTCTCATCTCTGGACA--CTTGTCTCTCGGTAGGTGGATGATTT 4293  
 3005 ACTTTTTRGCGCYRVTTCAGAAACCTTGTGCAATCAAGCCACCCMAGCRCTCTTMAATTT 3064  
 4294 ACTTTTGGCGCCCATTCAGAAACCTTGTGCAATCAAGCCACCCMAGCRCTCTTCAATTT 4353  
 3065 CTTGCTACCTGTGCTACAGGTTTCCAAACACCAAGGCTCAACTCTCTCTCACACAGGT 3124  
 4354 CTTGCTACCTGTGCTACAGGTTTCCAAACACCAAGGCTCAACTCTCTCTCACACAGGT 4413  
 3125 TAAATACCTTAGGCTTAARATTTATCCAAAGGCACCAAGGCCCTCAGTAGGGAAYRATCCA 3184  
 4414 ----TACTTAGGCTTAARATTTATCCAAAGGCACCAAGGCCCTCAGTAGGGAACACATCCA 4469  
 3185 GCCTATACCTGCTTATCTCATCTCAAAACCTTAAAGCACTAAGRRRTTCTCTTGGGT 3244  
 4470 GCCTATACCTGCTTATCTCATCTCAAAACCTTAAAGCACTAAGGGGATTTCTCTTGGGT 4529  
 3245 AAYAGGTTCTGCCGAATATGATTTCCCAAGTWTGGCAATAGCCAGGTCATTAWATA 3304  
 4530 AATAGGTTCTGCCGAATATGATTTCCCAAGTWTGGCAATAGCCAGGTCATTAWATA 4588  
 3305 CASTAATTAAGGAACCTCAGAAAGCAATACCCATTTARTAAGATGGAYAMCTGAAGYMR 3364  
 4589 CACTAATTAAGGAACCTCAGAAAGCAATACCCATTTAGTAGATGGACACTGAAGTAG 4648  
 3365 AAGTGGCTTTCCAGGCCCTTAAAGAGGCCCTTAAACCCAGYCCAGTGTGTTAAGYTTGCC 3424  
 4649 AAGTGGCTTTCCAGGCCCTTAAACCCAGYCCAGTGTGTTAAGYTTGCC 4694  
 3425 AACGGGCAAGACTTTTATATATTCACAGAAACCAAGAAAGTCTCTGAGGTCC 3484  
 4695 AACGGGCAAGACTTTTATATATTCACAG-AAAAAACAGGAATAGCTCTGAGGATCC 4753  
 3485 TTACACAGTCTCAGGGGAGTGTGCAACCVRTGCRVACCTGASTAAGGAAATGATG 3544  
 4754 TTACACAGTCTCAGGGGAGTGTGCAACCTGTGGCATACCTGACTAAGGAATGATG 4813  
 3545 TAGTGGCAAAAGGTTGRCVTCATTTTAYGGGTAGTGGTGCAGTGCAGYKTAGTAT 3604  
 4814 TAGTGGCAAAAGGTTGACCTCTATTGTTACGGGTAGTGGTGCAGTGCAGTCTTAGTAT 4873  
 3605 CTGAGCAGTATAAATAATACAGGAGAGATCTTACTGTGGGACATCTCATGAGTCA 3664  
 4874 CTGAGCAGTATAAATAATACAGGAGAGATCTTACTGTGGGACATCTCATGAGTCA 4933  
 3665 AYRGCACTACTCACTCTAAAGGAGACTTGTGGCTGTGAGCAACATGTTTACTTAAATRTC 3724  
 ::::



4934	Db	ATGGCATACTCAGCTCTAAAGGAGACTTGTGGCTGTGACAACTGTTTAACTTAAATGTC	4999
3725	Qy	AGGCTCTATTACTTTGAARGGCCAGTGTGCTGCRACCTGTGCACTGTGTGCAACTCTTAACCCAG	3784
4994	Db	AGGCTCTATTACTTTGAARGGCCAGTGTGCTGCGACTGTGCACTGTGTGCAACTCTTAACCCAG	5053
3785	Qy	YCNCAATTTCTTCCAGACAATCAAGAAAGAGATARAAYATPACTGTCAACAARTAAATTTCTC	3844
5054	Db	CCACATTTCTTCCAGACAATGAAGAAAGATATAAACATAACTGTGTCAACAAGTAATTTCTC	5113
3845	Qy	AAACCTATGCCACTCGAGGGGACCTTGTAGARGTTCCVTTTGACTGATCCYGCCTTCAAC	3904
5114	Db	AAACCTATGCCACTCGAGGGGACCTTGTAGAGTTCTTTGACTGATCCCGACC-TCAAC	5172
3905	Qy	TTGTATACTGATGAAGTTCTCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGCAGTG	3964
5173	Db	TTGTATACTGATGAAGTTCTTTGTAGAAAAAGGACTTCGAAAAGTGGGGTATGCAGTG	5232
3965	Qy	GTCAGTGATATGCAATATTGAAAGTAATCCCTCTACTCCAGAACTAGTGTCTYAGCTR	4024
5233	Db	GTCAGTGATATGCAATATTGAAAGTAATCCCTCTACTCCAGAACTAGTGTCTYAGCTR	5292
4025	Qy	GCAGAACTAATAGCCYCTCAVTKGGCCACTAGAAATTTAGGAGAAGRAAAAAGGVAATATA	4084
5293	Db	GCAGAACTAATAGCCCTCACTTGGGCCTAGNAATTAGAGAGAGAAAAGGCAATATA	5352
4085	Qy	TATACAGACTCTRATATGTCTYACCTAGTCTNCTCATGCCATCMRGCAATATATGSARAGAA	4144
5353	Db	TATACAGACTCTAAATATGCTTACTAGTCTCTCCATGCCCATATATGGAAGAA	5412
4145	Qy	AGGCAATTCCTAACTTCYAGRGAACHACTTATCMACATCAGGAAGCCATTAGABARATTA	4204
5413	Db	AGGCAATTCCTAACTTCCTGAGAGAACHACTTATCAACATCAGGAAGCCATTAGAGAAATTA	5472
4205	Qy	TTATVGGCWGTACAGAAACCTTARAGAGTGTGAGTCTTTACACTCTCTGGGGTCTATCANAAA	4264
5473	Db	TTATVGGCTGTACAGAAACCTTAAAGAGTGTGAGTCTTTACACTCTGGGGTCTATCAGAAA	5532
4265	Qy	GGAAAGRAAAGGGAAATASAGRGAAATYTGCAAGCAKATATTGAAGCMAAAGAGCTGCA	4324
5533	Db	GGAAAGRAAAGGGAAATAGAAAGAACTGCCNAGCAGATATTGAAGCCAAAAGAGCTGCA	5592
4325	Qy	AGGCAGAGCCCTCATTTAGAAAATGCTTTATTAACCTTCCCTTAGTATAGGGTAATCCCTTC	4384
5593	Db	AGGCAGAGCCCTCATTTAGAAAATGCTTTATTAATAACACCCCTAGTATAGGGTAATCCCTTC	5652
4385	Qy	CGGGAACCAAGCCCACTACTCAGCAGGAGAAACAGAAATGGGGAACCTCACAGG-CAG	4443
5653	Db	CGGGAACCAAGCCCACTACTCAGCAGGAGAAACAGAAATGGGGAACCTCACAGAGACAG	5712
4444	Qy	TTTTTCTCCCTCTGGGACGGTTAGCCACTGAAGAGGGGAAAATACTTTTGCTCTCAACTAT	4503
5713	Db	TTTTTCTCCCTCTGGGACGGTTAGCCACTGAAGAGGGGAAAATACTTTTGCTCTCAACTAT	5772
4504	Qy	CCAATGGAATTTACTTTAAAACCCCTTCATCAAACTTTCACTTAGGCATCGATAGCACCCA	4563
5773	Db	CCAATGGAATTTACTTTAAAACCCCTTCATCAAACTTTCACTTAGGCATCGATAGCACCCA	5832
4564	Qy	TCBARTGGCCAAATCATTTATTACTGCAACGAGCTTTTCAAACTATCAGAGATATXKT	4623
5833	Db	TCBARTGGCCAAATCATTTATTACTGCAACGAGCTTTTCAAACTATCAGAGATATAGT	5892
4624	Qy	CAGGGCTGTGAATGTGCCARAAAATAATCCCTGCCCTYATCGCCAAGCTCTTTCAGG	4683
5893	Db	CAGGGCTGTGAATGTGCCARAAAATAATCCCTGCCCTYATCGCCAAGCTCTTTCAGG	5952
4684	Qy	ARAAACAARAACAGGCCATTTACCTGTGRARAARACTGTGCAACTGATTTTATCCCAAGCCC	4743
5953	Db	ARAAACAARAACAGGCCATTTACCTGTGGAAGAAGCTGGCAACTGATTTTATCCCAAGCCC	6012
4744	Qy	AAACCTCAGGGAATTTCAGTATCTATAGTCTGGGTARATCTTTCACGGGTTGGGCARAG	4803
6013	Db	AAACCTCAGGGAATTTCAGTATCTATAGTCTGGGTAGATCTTTTCACGGGTTGGGCAGAG	6072

[illegible]

```
QY 5884 GTCACTGTCTGTGGACTTACTTACCACCAAACTGGTATGTCTGTATGGGGGTGGAGTTCAA 5943
Db |||
QY 7152 GTCACTGTCTGTGGACTTACTTACCACCAAACTGGTATGTCTGTATGGGGGTGGAGTTCAA 7211
Db |||
QY 5944 GATCAGGCAAGAGAAAACATGTAAAGAAAGTAATCTCCCAACTCACCGGGGTATCATGGC 6003
Db |||
QY 7212 GATCAGGCAAGAGAAAACATGTAAAGAAAGTAATCTCCCAACTCACCGGGGTATCATGGC 7271
Db |||
QY 6004 ACTCTAGCCCTTACAAGAGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCCAT 6063
Db |||
QY 7272 ACTCTAGCCCTTACAAGAGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCCAT 7331
Db |||
QY 6064 ACTCCGCTGTGAAGCTATTTAATACCAACCTCACTGGGCTCCATGAGGCTCTCGGCCCAA 6123
Db |||
QY 7332 ACTCGCCTGGTAAGCTATTTAATACCAACCTCACTGGGCTCCATGAGGCTCTCGGCCCAA 7391
Db |||
QY 6124 AACCCCTACTAAGCTGTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATCCCT 6183
Db |||
QY 7392 AACCCCTACTAAGCTGTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATCCCT 7451
Db |||
QY 6184 GTACCTGAACAATGGAACAACTTACGACACAGAAATAAACAACACACTTCCTGTTTTAGTAGGA 6243
Db |||
QY 7452 GTACCTGAACAATGGAACAACTTACGACACAGAAATAAACAACACACTTCCTGTTTTAGTAGGA 7511
Db |||
QY 6244 CCTCTGTGTTCCAAATSTGGAAATACCCATACCTCAAACTCACCTGTGTRAAATTTAGC 6303
Db |||
QY 7512 CCTCTGTGTTCCAAATSTGGAAATACCCATACCTCAAACTCACCTGTGTRAAATTTAGC 7571
Db |||
QY 6304 AATACTACATACACAACCAACTCCCAATGCATCAGTGGGTAACTCCTCCCAACACAAATA 6363
Db |||
QY 7572 AATACTACATACACAACCAACTCCCAATGCATCAGTGGGTAACTCCTCCCAACACAAATA 7631
Db |||
QY 6364 GTCTGCTACCTCAGGAATATTTTTGTGTGTGTACCTCAGCCCTATGTTGTTTGAAT 6423
Db |||
QY 7632 GTCTGCTACCTCAGGAATATTTTTGTGTGTGTACCTCAGCCCTATGTTGTTTGAAT 7691
Db |||
QY 6424 GGCTCTTCAGAACTATATGTGCTTCCTCTCATTTCTTAGTGCCCCCTATGRCATCTACACT 6483
Db |||
QY 7692 GGCTCTTCAGAACTATATGTGCTTCCTCTCATTTCTTAGTGCCCCCTATGACCACTACACT 7751
Db |||
QY 6484 GAACAAGATTTATACAGTTATGTCTATCTAAGCCCGCAACAAAGAGTACCCATTTCT 6543
Db |||
QY 7752 GAACAAGATTTATACAGTTATGTCTATCTAAGCCCGCAACAAAGAGTACCCATTTCT 7811
Db |||
QY 6544 CTTTTGTTTATAGGACGAGAGTGTCTAGTGCACTAGTGTACTGGCAATTTGGCGGTATCACA 6603
Db |||
QY 7812 CTTTTGTTTATAGGACGAGAGTGTCTAGTGCACTAGTGTACTGGCAATTTGGCGGTATCACA 7871
Db |||
QY 6604 ACTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAAACGGGTC 6663
Db |||
QY 7872 ACTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAAACGGGTC 7931
Db |||
QY 6664 GCGACTCCTCGTCACTTGTCAAGATCAACTTAACTCCTTAGCAGCAGTAGTCTTCRA 6723
Db |||
QY 7932 GCGACTCCTCGTCACTTGTCAAGATCAACTTAACTCCTTAGCAGCAGTAGTCTTCRA 7991
Db |||
QY 6724 AATCGAAGAGCTTTAGACTTGTCTAAACCGCTGARAGAGGGGAACTGTTTATTTTAGGG 6783
Db |||
QY 7992 AATCGAAGAGCTTTAGACTTGTCTAAACCGCTGARAGAGGGGAACTGTTTATTTTAGGG 8051
Db |||
QY 6784 GAAGATGCTGTTATTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 6843
Db |||
QY 8052 GAAGATGCTGTTATTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 8111
Db |||
QY 6844 CSAGATCGAATAACAAGTAKAGCAGAGAGCTTTCGAAACACTGGACCTCGGGCCCTCCCTC 6903
Db |||
QY 8112 CGAGATCGAATAACAAGTAKAGCAGAGAGCTTTCGAAACACTGGACCTCGGGCCCTCCCTC 8171
Db |||
QY 6904 AGCCATGATGATGCTTGGATTTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATATTTGCTA 6963
Db |||
QY 8172 AGCCATGATGATGCTTGGATTTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATATTTGCTA 8231
Db |||
QY 6964 CTCCTCTTTGGACCCCTGTATCTTTTACCTCCTTGTGTTAACTTTGTCTCTCCAGAAATCGAA 7023
Db |||
```

```
Db |||
QY 8232 CTCCTCTTTGGACCCCTGTATCTTTAACTCTCTTGTAACTTTGTTCTCTTCAGAAATCGAA 8291
Db |||
QY 7024 GCTGTAAACTTACAAATGGAGCCCAAGATGCAGTCCAAAGACTAAAGACTACCGCAGACCC 7083
Db |||
QY 8292 GCTGTAAACTTACAAATGGAGCCCAAGATGCAGTCCAAAGACTAAAGACTACCGCAGACCC 8351
Db |||
QY 7084 CTGACCGGCTGTGTAGCCACGATCTGTATGTTTAAATGACATCAAAAGGACCCCTCCTGAG 7143
Db |||
QY 8352 CTGACCGGCTGTGTAGCCACGATCTGTATGTTTAAATGACATCAAAAGGACCCCTCCTGAG 8411
Db |||
QY 7144 GAAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 7203
Db |||
QY 8412 GAAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 8471
Db |||
QY 7204 SGTGGCCCAACTCCCAACAGCACTTAGTGTTCCTGTTGAGATGGGGGAC 7255
Db |||
QY 8472 CGTGGCCCAACTCCCAACAGCACTTAGTGTTCCTGTTGAGATGGGGGAC 8523
Db |||
```

## RESULT 8

```
US-10-450-763-20014
; Sequence 20014, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 20014
; LENGTH: 6394
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (4994)..(6322)
; OTHER INFORMATION: 82% homologous to multiple sclerosis associated retrovirus
; OTHER INFORMATION: polyprotein, accession number AF009668, Smith-Waterman Score=1859.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(6394)
; OTHER INFORMATION: n = a,t,c or g
US-10-450-763-20014
```

```
Query Match 49.9%; Score 3738.4; DB 10; Length 6394;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 4026; Conservative 146; Mismatches 190; Indels 55; Gaps 21;
```

```
QY 1 CAACAATCGGATATAAACCCAGGCAATTCGAGCTGGCAACAGCAGCCCTTTGGGTCC 60
Db |||
QY 1959 CAAGGATGCAATATAAACCCAGGCAATTCGAGCAGCAACGCGCAACCGCTTTGGGTCC 2018
Db |||
QY 61 CTTCCCTTTGTATGGAGCTGT-----TTTCATGCTATTCACTCTATTAAT 108
Db |||
QY 2019 CTTCCCTTTGTATGGAGCTGT-----TTTCATGCTATTCACTCTATTAAT 2078
Db |||
QY 109 CTTGCAACTG--CACTCTTCTGCTCCATGTTCTTACGCTCGAGCTGAGCTTTTGCTCA 166
Db |||
QY 2079 CTTGCAACTGACACTCTCTGTCAGTGTGTTTACAGCTCTAGCTGAGCTTTGCTTG 2138
Db |||
QY 167 CGGTCCACCACTGCTGTTT---GCCACCAACCGCANACCTGCGCTGACTCCCATCCCTCT 223
Db |||
QY 2139 CAGTCCACCACTGCTGTTTGGCGCGCCACCGCAGACCGGCTGCTGCTTCCATCCCTTC 2198
Db |||
```

QY 224 GGATCTCTGAGGGTGTCCGTCTCTGATCCAGCGARGCGCCCATTTGCGCTCCCAA 283  
DB 2199 GGATCCGCGAGGGTGTCTGTGATTCCTGATCCAGCGAGGGCCCATTTGCGCGCCCAA 2258  
QY 284 TTGGGCTAAAGGCTTGCCATGTGNCCTGCAAGGCTAAGTCCCTGGGTTTGTCTTAATTGA 343  
DB 2259 TCGGGCTAAAGGCTTGCCATTTGCTCTGCAAGGTTAAGTCCCGGGTTGGTTCGTAATCGA 2318  
QY 344 GCTGACACTANTACACTGCGGTTCATGGTTCTCTCTGACCCCAAGGCTTCTAATTA-KA 402  
DB 2319 GCTGAAACACTTGTCACTGGGTTCACGGTTCTCTTCCGTGACCCCATGGCTTCTAATAGGA 2378  
QY 403 ACTAATAACCTTACCACAT-GGGCCAAAGATTCCATTCCTT-GGAATCCGTGAGGSCAAAG 460  
DB 2379 ACTAATAACATTACCACATGGGCCCAAGATTCCATTCCTTGGGAATTCGCTGAGGCCAAG 2438  
QY 461 AACTCCAGGTCAGAGATAAGAGCTTGGCCACCATTCTGGAAGCGGCTCTGCTACCTTCCT 520  
DB 2439 AACCCAGGTCAGAGATAAGGCTTGGCCACCATTCTGGAAGCGGCTCTGCTACCTTCCT 2498  
QY 521 GGAAGTGGTTACCAACATCTTGGGAGCTCTGTGAGCAAGGACCCCGGCTGACATTTTG 580  
DB 2499 GGAAGCGGCTTGGCCACATCTTGGGAGCTCTGTGCAACAGGACCCCGGCTAATTTG 2558  
QY 581 GCRACAMSRACGACATCCMAAGTGTAGTGGAAACGTTCCCGCAGACAAAACGCCCC 640  
DB 2559 GCAACACGACGAGCATCCAAAGTGTAGTGGAAACGTTCCCGCAGACAAAACGCCCC 2618  
QY 641 TAAGAGCTATTCTGGARAATTGGGAMCAATTTGACCCCTCAGACACTAA-GAAAGAAACGA 699  
DB 2619 TAAGAGCTATTCTGGAGAAATTGGGACCAATTGACCCCTCAGACACTAAGGAAGAAACGA 2678  
QY 700 CTTATATTCTTCTGAGTGCGGCTGGCACTCTGAGGGAAGTATAAATTTAATTAACACCAT 759  
DB 2679 CTTATATTCTTCTGAGTGCGGCTGGCACTCTGAGGGAAGTATAAATTTAATTAACACCAT 2738  
QY 760 CTTACAGCTAGACVCTCTTTTGTAGAA--RAGGCAATGAGTGAAGTCCCAAGTACA 816  
DB 2739 CTTACAGCTAGACTCTTTTGTAGAAAGAGTAAATGGAGTGAAGTGTATATGTACA 2798  
QY 817 AACTTTCTTTTCAATTAAGAGACAACCTACAAATTA-TGTAAAAAGTGTGATTTATGCC-- 873  
DB 2799 AACTTTCTTTTCAATTAAGAGACAACCTACAAATTTATTTGTAAAGTGTGATTTATGCCCT 2858  
QY 874 TACAGGAAGCTTTCAGAGTCTACCTCCCTATCCC--AGCATCCCGCACTCCTTCCCCA 929  
DB 2859 ACAGGGAAGCTTTCAGAGTCTACCTCCCTATCCCAGCATCCCCCACTCCTTCTCCA 2918  
QY 930 MYTAATAAGACCCCTTCAACCCAAATGGTCCAAAGAGAGATAGACAAAAGGGTAAAC 989  
DB 2919 ACTATTAAAGGACCCCTTCAACCCAAATGGTCCAAAGAGAGATAGACAAAAGGGTAAAC 2978  
QY 990 AGTGAACCAAGAGTGCCCAATATTTCCCAATATGACCCCTCCCAAGCATGTGGAGGAAG 1049  
DB 2979 AGTGAACCAAGAGTGCCCAATATTTCCCAATATGACCCCTCCCAAGCATGTGGAGGAAG 3038  
QY 1050 AGAATTCGGGCCACGAGAGTGATGTGCTTTTCTCCAGACTTAAAGCAATAAA 1109  
DB 3039 AGAATTCGGGCCACGAGAGTGATGTGCTTTTCTCTCCAGACTTAAAGCAATAAA 3098  
QY 1110 AACAGACTTAGGTAAATTTTCAGATAAYCCTGATGCTATATTTGRTGTTTTACAGGGTT 1169  
DB 3099 AACAGACTTAGGTAAATTTTCAGATAAACCTGATGCTATATTTGGTGTTTTACAGGGTT 3158  
QY 1170 AGGACAAATTTTGAATTCAGATGGAGAGATATATATGTCATGCTGCTAATTCAGACACTAA 1229  
DB 3159 AGGACAAATTTTGAATTCAGATGGAGAGATATATATGTCATGCTGCTAATTCAGACACTAA 3218  
QY 1230 CCCCAAATGAGAAAGTGCACCATAACTGCAGCTGAGGTTTGGCGATCTCTGCTATC 1289  
DB 3219 CCCCAAATGAGAAAGTGCACCATAACTGCAGCTGAGGTTTGGCGATCTCTGCTATC 3278  
QY 1290 TCAGTCAGGTCAATGGATANGGATGACAAAGAGAAAGANAAATGATTTCCCAAGGCC 1349

DB 3279 TCAGTCAGGTCAATGGATA--GGATGACACAG-AGAAAAGAGACGATTTCCCAACAGGCC 3336  
QY 1350 AGCARGCAGTTCACAGTCTASACCTCATTTGGGGACACAGAAATCAGTAAACATCGGAGAT 1409  
DB 3337 AGCARGCAGTTCACAGTCTAGACCTCATTTGGGGACACAGAAATCAGTAAACATCGGAGAT 3396  
QY 1410 TGTGCTGACAGACATTTGCTAACTTGTGCTCTASAGGACTAAGGAAAACTASGAGAAAA 1469  
DB 3397 TGTGCTGACAGACATTTGCTAACTTGTGCTCTCAAGGACTAAGGAAAACTACGAGAAAA 3456  
QY 1470 RTCTATGAATTTACTCAATGATGTCACCAATAACAGGGGAGGAGAAATTCCTACTG 1529  
DB 3457 ATCTAGAAATTTACTCAATGATGTCACCAATAACAGGGGAGGAGAAATTCCTACTG 3516  
QY 1530 CTTTCTGAGAGACTAAGGGAGGCAATTGAGGAAGGCTGCTCTCTGCTACCTGACTCTT 1589  
DB 3517 CTTTCTGAGAGACTAAGGGAGGCAATTGAGGAAGGCTGCTCTCTGCTACCTGACTCTT 3576  
QY 1590 CTGAAGGCCAACTAATCTTTAAAGCGTAAGTTTATCACTCAGTGCAGACATTAAGAA 1649  
DB 3577 CTGAAGGCCAACTAATCTTTAAAGCGTAAGTTTATCACTCAGTGCAGACATTAAGAA 3636  
QY 1650 AAAAATTTCAAAAAGTCTGCCGTAGGCCGAGGCAAACTTAGAAAACTTGAATTTGGC 1709  
DB 3637 AAAAATTTCAAAAAGTCTGCCGTAGGCCGAGGCAAACTTAGAAAACTTGAATTTGGC 3696  
QY 1710 AACVTCGGTTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAAACAGACAAACGGGATTA 1769  
DB 3697 AACVTCGGTTTTTTTATAATAGAGATCAGGAGGAGCAGGCGGAAACAGACAAACGGGATTA 3756  
QY 1770 AAAAAAGGCCAACCGCTTTAGTCATGACCCCTCAGGCAAGTGAGACTTTTGGAGGCTCTGAA 1829  
DB 3757 AAAAAAGGCCAACCGCTTTAGTCATGACCCCTCAGGCAAGTGAGACTTTTGGAGGCTCTGAA 3816  
QY 1830 AAGGGAAGAGCTGGGCAAAATTTGAATGCTTAAAGGCTTGTCTCAGTGGGCTCTACAAG 1889  
DB 3817 AAGGGAAGAGCTGGGCAAAATTTGAATGCTTAAAGGCTTGTCTTCCAGTGGGCTCTACAAG 3876  
QY 1890 GACACTTTAAAAAAGATTTGTCAAAGTAGAAGTAAGCGCCCTTCGTCCATGCGCCCTTAT 1949  
DB 3877 GACACTTTAAAAAAGATTTGTCAAAGTAGAAGTAAGCGCCCTTCGTCCATGCGCCCTTAT 3936  
QY 1950 TTCAAGGGAATCACTGGAAGGCCCACTGCCCAAGGGGACAAAGGTCTTTTGAAGTCAGAAG 2009  
DB 3937 TTCAAGGGAATCACTGGAAGGCCCACTGCCCAAGGGGACAAAGGTCTTTTGAAGTCAGAAG 3996  
QY 2010 CCACATAACAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGGCCATCCCATG 2069  
DB 3997 CCACATAACAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGCAAGGCCATCCCATG 4056  
QY 2070 CCATCACCTTCAGAGCCCTGGGTATGCTTGCACCAATTCAGGGCCAGGAGGTGTCTCC 2129  
DB 4057 CCATCACCTTCAGAGCCCTGGGTATGCTTGCACCAATTCAGGGCCAGGAGGTGTCTCC 4116  
QY 2130 TGGACACTGGTGGGCTTCTTTAGTCTTACTCTTCTGTCGGGACAACTGTCTCCAGAT 2189  
DB 4117 TGGACACTGGTGGGCTTCTTTAGTCTTACTCTTCTGTCGGGACAACTGTCTCCAGAT 4176  
QY 2190 CTGTCACTATTCTGAGGGGCTCCNTAAGAAGGCGAGTCACTAGATACTTTTCCAGGCCA 2249  
DB 4177 CTGTCACTATTCTGAGGGGCTCC-TAAGACGGGAGTCACTAGATACTTTTCCAGGCCA 4235  
QY 2250 CTAAGTTATGAAGTGGGAGCTTTTATCTTTTCACTGCTTTTCAATTTATGCTGAAAG 2309  
DB 4236 CTAAGTTATGAAGTGGGAGCTTTTATCTTTTCACTGCTTTTCAATTTATGCTGAAAG 4295  
QY 2310 CCCCCTACTCTTTAGGGAGAGACATTTCTAGCAAAAGCAGGGCCATTTATACCTGAA 2369  
DB 4296 CCCCCTACTCTTTAGGGAGAGACATTTCTAGCAAAAGCAGGGCCATTTATACCTGAA 4355  
QY 2370 CATAGGAGAGGAACACCCGTTTGTGTCCTGCTTGGAGGAAGAAATTAATTCCTGAAAG 2429

Db 4356 CATAGGAGAGGAAACACCGTTTGTGT - CCCTGCTTGAGGAAGAAATTAATCCTGAAG 4414  
QY 2430 TCTGGGCAACAGAGGCAATATGACGAGCCAAAGATGCCGCTCCTGTTCAAGTTAAA 2489  
Db 4415 TCTGGGCAACAGAGGCAATATGACGAGCCAAAGATGCCGCTCCTGTTCAAGTTAAA 4474  
QY 2490 CTAAAGGATTCACATTCCTTCCCTTACCAAGGCAAGTACCCCTCAGACCCCAAGGCCAA 2549  
Db 4475 CTAAAGGATTCCTGCTCCTTCCCTTACCAAGGCAATACCCCTTAGACCTGAGGCCCA 4534  
QY 2550 CAA - GGAATCCAAAGATGTTAAGGACTTAAAGCCCAAGCTTAGTAAACCATGCA 2607  
Db 4535 CAAAGGCTTCCAAAGATGTTAAGGACTTAAAGCCCAAGCTTAGTAAACCATGCA 4594  
QY 2608 TAACTCCCTGCAAGTAAATCCGTTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACGTGG 2667  
Db 4595 TAACTCCCTGCAAGTAAATCCGTTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACGTGG 4654  
QY 2668 AGGGTTAGTGAAGATCTCAGGATTAATCAATGGAGCCGTTCTCTTTTATACCCAGCTG 2727  
Db 4655 AGGGTTAGTGAAGATCTCAGGATTAATCAATGGAGCCGTTCTCTTTTATACCCAGCTG 4714  
QY 2728 TACCTAGCCCTTATCTGMYTTCCTCAATACCAAGGAGCAGAGTGGTTTACASTCC 2787  
Db 4715 TACCTAGCCCTTATCTGCTGCTTCCCAATATACCAGAGCAGAGTGGTTTACASTCC 4774  
QY 2788 TGGACCTTMAAGATGCTTCTTGCAATCCCTGTACATCCTGACTCTCAATCTTGTGTTG 2847  
Db 4775 TGGACCTTMAAGATGCTTCTTGCAATCCCTGTACATCCTGACTCTCAATCTTGTGTTG 4834  
QY 2848 CTTTGAAGATCTTCAAAACCCARCATCTCACTCAGTGGACTTTTACCCCAAGGCT 2907  
Db 4835 CTTTGAAGATCTTCAAAACCCARCATCTCACTCAGTGGACTTTTACCCCAAGGCT 4894  
QY 2908 TCAGGATAGYCCCATCTATTTGGCCAGGCAATAGCCCAAGACTTGAGYCAATYMTAT 2967  
Db 4895 TCAGGATAGYCCCATCTATTTGGCCAGGCAATAGCCCAAGACTTGAGYCAATYMTAT 4954  
QY 2968 ACTTGACACTCTGCTTCCTGRTAKGTGGATG - ATTACTTTTTRGCGCCVRRCTCAGAA 3026  
Db 4955 ACTTGACACTCTGCTTCCTGRTAKGTGGATG - ATTACTTTTTRGCGCCVRRCTCAGAA 5014  
QY 3027 ACTTGTGCAATCAAGCCACCAAGCCTCTTAAATTTCTCGVACCGTGGCTACAGW 3086  
Db 5015 ACTTGTGCAATCAAGCCACCAAGCCTCTTAAATTTCTCGVACCGTGGCTACAGW 5074  
QY 3087 GTTTCACAAACSAARGCTCACTCTGCTCACAGCAGGTTAAATCTTAGGCTAAATTA 3146  
Db 5075 GTTTCACAAACSAARGCTCACTCTGCTCACAGCAGGTTAAATCTTAGGCTAAATTA 5134  
QY 3147 TCCAAAGGACCAAGCCCTCAGTGAGGAAAYATCCAGCCTATATCTGGCTTATCCTCAT 3206  
Db 5135 TCCAAAGGACCAAGCCCTCAGTGAGGAAAYATCCAGCCTATATCTGGCTTATCCTCAT 5194  
QY 3207 CYCAAAACCTTAAGCACTAAGRPRTTCTTGGCTTAAYAGGVTCTGCCGAATGG 3266  
Db 5195 CCCAAACCTTAAGCACTAAGRPRTTCTTGGCTTAAYAGGVTCTGCCGAATGG 5254  
QY 3267 ATTCCCAAGTGTGGCAATATAGCCAGGCTTAAATATACATACTAATTAAGGAACTCAGAA 3326  
Db 5255 ATTAA - CAGTATGGCAATATAGCCAGGCTTAAATATACATACTAATTAAGGAACTCAGAA 5313  
QY 3327 AGCCAAATACCCATTTARTAAGATGGAYAMCTGAAGYMAAGTGGCTTTCCAGGCCCTTAA 3386  
Db 5314 AGCCAAATACCCATTTARTAAGATGGAYAMCTGAAGYMAAGTGGCTTTCCAGGCCCTTAA 5369  
QY 3387 AGAAGCCCTTAAACCAAGYCCAGTGTAAAGTTCGCAACGCGGCAAGACTTTSTTVA 3446  
Db 5370 -----TAAACCAAGCCCGTGTAAAGTTCGCAACGCGGCAAGACTTTSTTVA 5419  
QY 3447 TAYRTACAGAAAAACAGAAAYAGCTCTGAGGTCCTTACACAGRTCCAGGGAGYAG 3506  
Db 5420 TATGTCACAGAAA - AAACAGGAATAGCTCTAGGAGTCTTACACAGATCCGAGGATGAG 5478

QY 3507 CTTGCAACCYRTGGCRYACCTGASTAAGAAAYTGTATGTAGTGGCAAGGGTTGRCYTCA 3566  
Db 5479 CTTGCAACCTGTGGGTACCTGACTAGTGAAGAAATTTAGTGGCAAGGGTTGRCYTCA 5538  
QY 3567 TTGTTTATGGGTAGTGGTGGCAGTAGCAGTGYKTAGTATCTGAAGCAGTTAAATAATACA 3626  
Db 5539 TTGTTTATGGGTAGTGGTGGCAGTAGCAGTGYKTAGTATCTGAAGCAGTTAAATAATACA 5598  
QY 3627 GGRAGAGATCTTACTGTGTGGAATCTCATGAKGTGAAYRGCACTACATCTGCTTAAAGG 3686  
Db 5599 GGAAGAGATCTTACTGTGTGGAATCTCATGAKGTGAAYRGCACTACATCTGCTTAAAGG 5658  
QY 3687 AGACTTGTGGCTGTGCAGAAACGTTTACTTAAATCTCAGGCTCTTACTTGAAGGCTC 3746  
Db 5659 AGACTTGTGGCTGTGCAGAAACGTTTACTTAAATCTCAGGCTCTTACTTGAAGGCTC 5718  
QY 3747 AGTGTGCTGCTGTGCACCTTGTGCAACTCTTAAACCCAGYCNCAATTTCTTCCAGACAATGA 3806  
Db 5719 AGTGTGCTGCTGTGCACCTTGTGCAACTCTTAAACCCAGYCNCAATTTCTTCCAGACAATGA 5778  
QY 3807 AGAAAGATARAAYATACTGTCAACAARTAAATTTCTCAAACTATGCCATCTCGAGGGA 3866  
Db 5779 AAAAAGATAGAACGTAACCTGTCAACAGGTGATTGCTCAAACTATGCCATCTCGAGGGA 5838  
QY 3867 CCTTGTAGAGCTTCCCTTGTGATCCGCTTCACTTCACTTGTATCTGATGGAAGTCTCT 3926  
Db 5839 CCTTGTAGAGCTTCCCTTGTGATCCGCTTCACTTGTATCTGATGGAAGTCTCT 5897  
QY 3927 TTGTAGAAAAAGGACTTCGAAAAAGYGGGCTATGCACTGCTGATGATGGAATATTTG 3986  
Db 5898 TTGTAGAAAAAGGACTTCGAAAAAGYGGGCTATGCACTGCTGATGGAATATTTG 5957  
QY 3987 AAAGTAATCCCTCACTCCAGAACTAGTGTAGTGTGCAAGAACTAATAGCCYCTATK 4046  
Db 5958 AAAGTAATCCCTCACTCCAGAACTAGTGTAGTGTGCAAGAACTAATAGCCYCTATK 6017  
QY 4047 GGGCACTAGAAATAGGAGAGGAAAGGAAATATATATATACAGACTCTTATATGCTY 4106  
Db 6018 GGGCACTAGAAATAGGAGAGGAAAGGAAATATATATATACAGACTCTTATATGCTY 6077  
QY 4107 ACCTAGTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 4166  
Db 6078 ACCTAGTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 6137  
QY 4167 GAACACCTATCAMACATCAGGAGCCATTAGGAAATTTATYTGCGWTGACAGAACCTA 4226  
Db 6138 GAACACCTATCAMACATCAGGAGCCATTAGGAAATTTATYTGCGWTGACAGAACCTA 6197  
QY 4227 RAGAGTGTGMACTTCTTACACTGCGGGGTCTCATANAAGGAAAGGAAATASAG 4286  
Db 6198 RAGAGTGTGMACTTCTTACACTGCGGGGTCTCATANAAGGAAAGGAAATASAG 6257  
QY 4287 RGAAYTGCAAGCAKATATTGAAGGMAAAGAGCTGCAAGGCGAGCCCTCCATTAGAAA 4346  
Db 6258 RGAAYTGCAAGCAKATATTGAAGGMAAAGAGCTGCAAGGCGAGCCCTCCATTAGAAA 6317  
QY 4347 TGCTTATTAAACTTCCCTTAGTATAGGGAATCCCTT 4383  
Db 6318 TGCTTATTAGAGGACCCCTTAGTATGGCAAGCAATT 6354

## RESULT 9

US-10-763-12279  
; Sequence 12279, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hysq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631

```
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 12279
; LENGTH: 4349
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (466)..(906)
; OTHER INFORMATION: 100% homologous to Homo sapiens neurotensin
; OTHER INFORMATION: receptor, accession number X70070, Smith-Waterman Score=775.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(4349)
; OTHER INFORMATION: n = a, t, c or g
US-10-450-763-12279

Query Match      45.8%; Score 3431; DB 10; Length 4349;
Best Local Similarity 87.3%; Pred. No. 0;
Matches 3974; Conservative 148; Mismatches 195; Indels 237; Gaps 28;

QY 115 ACTGCACCTCTTCGTGTCATGTTCTTACGGCTCGAGCTGAGCTTTTGTCTCACCGTCCAC 174
DB 1 ACTGCACCTCTTCGTGTCATGTTCTTACGGCTCGAGCTGAGCTTTTGTCTCACCGTCCAC 60

QY 175 CACTGCTGTTTGGCCACACCGCANACTGCGCGTGAATCCCATCCCTCTGGATCCTGCAG 234
DB 61 CACTGCTGTTTGGCCACACCGCANACTGCGCGTGAATCCCATCCCTCTGGATCCTGCAG 120

QY 235 GGTGTCGGCTGTCTCTGATCCAGGGA---RCGCGCCATTGCGGTCCTCCCAATTTGGGCTA 291
DB 121 GGTGTCGGCTGTCTCTGATCCAGGAGCGCGGCGCCATTGGCACTCCCGATCGGGGCTA 180

QY 292 AGGCTTGCCATTGTTNCTGTCAC-GGCTAAGTGCCT-GGGTTGTTCTAATTGAGCTGAA 349
DB 181 GAGGCTCACCATGTTTCTTGCAAGGCTAAGTGCCTGGGGTTTCATCTAATCGAGCTGAA 240

QY 350 CACTANTCACTGGTTCATGTTCTCTCTGACCCGCTGACCCGCTTCTAATAAATATAA 409
DB 241 CACTAATCACTGGTTCATGTTCTCTCTGACCCGCTGACCCGCTTCTAATAAATATAA 300

QY 410 CACTTACCATGCGCCCAAGATTCCATTCTCTGGAATCCGTGAGGSCAACGAACTCCAG 469
DB 301 CACTTACCATGCGCCCAAGATTCCATTCTCTGGAATCCGTGAGGSCAACGAACTCCAG 359

QY 470 TCAGAGAATACGAGCTTGCCACCATCTTTGGAAGCGGCTGTACCTCTTTGGAAGTGGT 529
DB 360 TCAGAGAATACGAGCTTGCCACCATCTTTGGAAGCGGCTGTACCTCTTTGGAAGTGGT 419

QY 530 TCACCAACCATCTTGGAGCTCTGTGACGAGGACCCCGGCTGACATTTTGGACACGMS 589
DB 420 TCACCAACCATCTTGGAGCTCTGTGACGAGGACCCCGGCTGACATTTTGGACACCAAC 479

QY 590 RACGGACATCCMAAGTGATGGGAAACGTTTCCCGCAAGACAAAACGCCCTTAAGACGTA 649
DB 480 GACGGACATCCMAAGTGATGGGAAACGTTTCCCGCAAGACAAAACGCCCTTAAGACGTA 539

QY 650 TTCTGGARAATTGGGAMCAATTTGACCCCTCAGACACTAAGAAAGAAACGACTTATATCT 709
DB 540 TTCTGGAGAATTGGGACCAATTTGACCCCTCAGACACTAAGAAAGAAACGACTTATATCT 599

QY 710 TCTGCAAGTCCGCTCGGCACTCTCTGAGGGAAGTATAAATAAACAACCTTTACAGCTA 769
DB 600 TCTGCAAGTCCGCTCGGCACTCTCTGAGGGAAGTATAAATAAACAACCTTTACAGCTA 659

QY 770 GACYTCTTTTGTAGAAAAGGCAATGGAGTGAAGTGCCTAAGTACAACTTTCTTTTCA 829
DB 660 GACCTCTTTTGTAGAAAAGGCAATGGAGTGAAGTGCCTAAGTACAACTTTCTTTTCA 719
```

```
QY 830 TTAAGAGACAACTCAAAATTTATCTAAAAAGTGTGATTTATGCCCTACAGGAAGCTTTCAG 889
DB 720 TTAAGAGACAACTCAAAATTTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCTTTCAG 779
QY 890 AGTCTACCTCCCTATCCCAAGCATCCCGACTCCTTCCCAATTAATAAGGACCCCTTTC 949
DB 780 AGTCTACCTCCCTATCCCAAGCATCCCGACTCCTTCCCAATTAATAAGGACCCCTTTC 839
QY 950 AACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAAACAAAGAGTGCAC 1009
DB 840 AACCCAAATGGTCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAAACAAAGAGTGCAC 899
QY 1010 TATTCCCAATTAAGACCCCTCCCAAGCATGGAGGAGAGAAATTCGGCCCAAGCCAGAG 1069
DB 900 TATTCCCAATTAAGACCCCTCCCAAGCATGGAGGAGAGAAATTCGGCCCAAGCCAGAG 958
QY 1070 TGCATGTGCVTTTTTCTTCCCAAGCATTAAGCAAAATTAATAAGAGCTTAGGTAAATCT 1129
DB 959 TGCATGTGCVTTTTTCTTCCCAAGCATTAAGCAAAATTAATAAGAGCTTAGGTAAATCT 1018
QY 1130 CAGATAAATCTGATGCTATATTGTTTCAAGGGTTAGGACAAATCTTTGATCTGA 1189
DB 1019 CAGATAAATCTGATGCTATATTGTTTCAAGGGTTAGGACAAATCTTTGATCTGA 1078
QY 1190 CATGGAGAGATATATGTCATGCTAAATCAGACACTAACCCCAATCAGAGAAAGTGCC 1249
DB 1079 CATGGAGAGATATA-ATGTCATGCTAAATCAGACACTAACCCCAATCAGAGAAAGTGCC 1137
QY 1250 ACCATAAATGTCACCTGAGGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATAN 1309
DB 1138 ACCATAAATGTCACCTGAGGTTTGGCG-TCTCTGGTATCTCAGTCAGGTCAATGGATAN 1196
QY 1310 GGATGACAAACAGAGAAAGAAATGATTTCCCAAGGCGCCAGCAGCTTCCAGCTTA 1369
DB 1197 GGATGACAAACAGAGAAAGAAATGATTTCCCAAGGCGCCAGCAGCTTCCAGCTTA 1256
QY 1370 SACCCTCATTTGGGGACACAGAAATCAGTAACATGGGAGATTTGGTGTGCAGACATTTGCT 1429
DB 1257 GACCCTCATTTGGG---ACACAGAAATCAGAACAT-GGAGATTTGGTGTGCAGACATTTGCT 1312
QY 1430 AACTTGTGTCTASAGGACTAAGGAAACCTASAGAAAGAAATCTATGAATTAATCAATGA 1489
DB 1313 AACTTGTGTCTASAGGACTAAGGAAACCTASAGAAAG-AAAGTCTATGAATTAATCAATGA 1371
QY 1490 TGTCCACCATAAACAGAGGGAAGGAAAGAAATCTTACTGCTTCTGAGGAGCTTAAGG 1549
DB 1372 TGTCCACCATAAACACA-GGGAAAGGGAAGAAATCTTACTGCTTCTGAGGAGCTTAAGG 1430
QY 1550 GAGGCATTGAGGAAGCGTGCCTCTCTGTCACTGACTCTTCTGAAGGCCAACTTAATCTTA 1609
DB 1431 GAGGCATTGAGGAAGCGTGCCTCTCTGTCACTGACTCTTCTGAAGGCCAACTTAATCTTA 1490
QY 1610 AAGCGTAAGTTTATCACTCAGTCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1669
DB 1491 AAGCGTAAGTTTATCACTCAGTCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1550
QY 1670 TAGGCGCGGAGCAAACTTAGAAACCTTATGAACTTGGCAACTCGGTTTTTATTAATA 1729
DB 1551 TAGGCGCGGAGCAAACTTAGAAACCTTATGAACTTGGCAACTCGGTTTTTATTAATA 1610
QY 1730 GAGATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1789
DB 1611 GAGATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1670
QY 1790 GTCTGAAGCTCAGCAAGTGGAGCTTTGGAGGCTCTGGAAGGAGGAGGAGGAGGAGGAGGAG 1849
DB 1671 GTCTGAAGCTCAGCAAGTGGAGCTTTGGAGGCTCTGGAAGGAGGAGGAGGAGGAGGAGGAG 1730
QY 1850 TGAATGCCCTAATAGGCTTGTTCAGTGGCGGTCTCAAGGACACTTTTAAAAAGATGTT 1909
DB 1731 TGAATGCCCTAATAGGCTTGTTCAGTGGCGGTCTCAAGGACACTTTTAAAAAGATGTT 1790
```



QY 1910 CCAAGTAGAAGTAAGCGGCCCTTCTGTCATGCCCTTATTTC AAGGAATCACTGGAAG 1969  
DB |||||  
DB 1791 CCAAGTAGAA----- 1800  
QY 1970 GCCCATGCCCCAGGGGCAAAAGGTCTTTTGTAGTCAGAAAGCACTAACAGATGATCCAG 2029  
DB |||||  
DB 1801----- 1800  
QY 2030 CAGCAGGACTGAGGGTGCCCTGGGGCAAGGCCATCCCATGCCATCACCTTCACAGAGCCC 2089  
DB |||||  
DB 1801-----AGCCC 1805  
QY 2090 TGGGTATGCTTCACCATTCAGAGGCGCAGGAAGGT---TGTCTCTCGACACTGTCGGTCT 2146  
DB |||||  
DB 1806 CAGGTATGCTTGACCATTTGAGAGCCAAAGAGGTAACTGTCTCGACCGCGGTGCGGCC 1865  
QY 2147 TTCTTAGTCTTACTCTCTGTCGCCGACAACTGTCTCCAGATCTGTCACTATCTTGAGG 2206  
DB |||||  
DB 1866 TTCTCAGTCTTACTCTCTGTCTCCGGAACAATGTCTCCAGATCTGTCACTATCC--GA 1922  
QY 2207 GGGTCNTAAGACGGGCGATCTAGATATCTTTTTCCTGAGCCACTAAGTTATGAATGGG 2266  
DB |||||  
DB 1923 GGGGCTCTAGGACAGCCACTCAGATAC--TTCTTCCAGCCACTAAGTTGTG-ACTGGG 1980  
QY 2267 GAGCTTTATTTTTCACATGCTTTTCTAATTAATGCTTGAAAGCCCACTACTCTGTTAG 2326  
DB |||||  
DB 1981 GAGCTTTTACCTTTTTCACATGCTTTCTAATTAATGCTTGAAAGCCCTCACTCCCTTGTAG 2040  
QY 2327 GGAGAGACATTTAGCAAAAGCAGGGGCCATTAATACACTGAACATAGAGAGGAAGAACAC 2386  
DB |||||  
DB 2041 AGAGAGAGATTCTAGTAAAGCAGGGGCCATTATACACTGAACATAGAGAGGAAGAACAC 2100  
QY 2387 CGGTTTGTGTCCTCTGTCAGGAGGAAGAAATTAATTCCTGAAGTCTGGGCAACAGAGGA 2446  
DB |||||  
DB 2101 CCAATTGTGCT--CTCCTGTTTGGAGGAAGAAATTAGTCTCTGAAGTCTGGGCAACAGAGGA 2159  
QY 2447 CAATAGGACGACCAAGAAAGTCCGCTCTGTCAGGTTAACTAAAGGATTCACATTC 2506  
DB |||||  
DB 2160 CAATATGGACAAG--CAAGAAATGCCCCATTTGTCTCAAGTTAACTAAAGGATTCACATTC 2218  
QY 2507 CTTTCCCTACCAAGGACGATACCCCTCAGACCCAGGCCCAACAGGATTCCAAAGAT 2566  
DB |||||  
DB 2219 CTTTCCCTATCAAGGACGATACCCCTCAGACCCGAGGCCCAACAGGATTCCAAAGAT 2278  
QY 2567 TGTTAAGGCTTAAAGCCAGGCTTAGTAAACCAATGCATAACTCCCTGCGAGTAAATTC 2626  
DB |||||  
DB 2279 TGTTAAGGACCTAAAGGCCAAGGCTTAGTAAACCAATACAGTAGCCCTGCAATCTCC 2338  
QY 2627 CGTAGTGGATTGAGGAGGCACAGAACCCAGTGGACAGTGGAGGTTAGTGCAGATCTC 2686  
DB |||||  
DB 2339 -----AATTTTAGGAGTACAGAAACCCAAAGACAGTGGGA--GGTTAGTGCAGATCTC 2390  
QY 2687 AGGATTATCAATGGAGCGGTGTCTCTTTTATACCCAGCTGTACCTAGGCCCTTATCTGT 2746  
DB |||||  
DB 2391 AGGATTATCAAT--GAGCTGTGTTTCTCTATACCCAGCTATACCTAGGCCCTTATCTCT 2449  
QY 2747 GMYTTCCCAATACCAAGGAAGCAGAGTGGTTTACATCTCTGACCTTGAAGATGCCCTT 2806  
DB |||||  
DB 2450 GCTTTCCTCCAAATACCAAGGAAGCAGTGGTTTACAGTCTCTGGATCTTAAAGGATGCCCTT 2509  
QY 2807 CTTCTGCATCCCTGTACATCTGACTCTCAATCTTGTGTTGCTTTGAAGATCTTCAAA 2866  
DB |||||  
DB 2510 TTTCTGCATCCCTGTACATCTGACTCTCAATCTTGTGTTGCTTTGAAGATCTTCAAA 2569  
QY 2867 CCAACATCTCAACTCACCTGGACTTTTATCCCAAGGGTTCCAGGATAGYCCCCATCT 2926  
DB |||||  
DB 2570 CCAACATCTCAACTCACCTGGACTTTTATCCCAAGGGTTCCAGGATAGTCCCATCT 2629  
QY 2927 ATTTGGCCAGGCAATTAGCCCAAGACTTGAGYCAATYMTATACCTGGAACACTCTGTCTCT 2986  
DB |||||  
DB 2630 ATTTGGCCAGGCAATTAGCCCAAGACTTGAGCCAGTTCTCATACCTGGAACACTCTGTCTCT 2689  
QY 2987 TCRGTAKGTGGATGATTTTACTTTTRCGYCCYRTTCAGAAACCTTGTGCTCAACGCCAC 3046

DB |||||  
DB 2690 TCAGTGCATGATGATTTACTTTTAGCTGCCGTTTCAGAAACCTTGTGCCATCAAGCCAC 2749  
QY 3047 CCAAGCRCTCTTTMAATTTCTCGCYACCTGTGGCTTACAWGGTTTCCAAACRSARGCTCA 3106  
DB |||||  
DB 2750 CCAAGCGCTCTTAAATTTCTCACTACCTGTGGCTTACAAGGTTTCCAAACCAAGGCTCA 2809  
QY 3107 RCTGCTCTCACAGCAGGTTAAATACTTAGGCTTAARATTTATCCAAAGCACCARGGCCCT 3166  
DB |||||  
DB 2810 GCTCTGCTCACAGCAGGTTAAATGCTTAGGGCTTAAATTTATCCAAAGTCCCHGGGCCCT 2869  
QY 3167 CAGTGAGGAAYRYATCCAGCCTTATCTGCTTATCTCATCYCAAAACCTTAAAGCAACT 3226  
DB |||||  
DB 2870 CAGTGAGGAACGTATCCAGCCTTATCTGCTTATCTTATCCCAAAACCTTAAAGCAACT 2929  
QY 3227 AAGRRRTTCTTGGCTTAAAYAGYTTCTGCGAATATGGAATTCCTCCAGGTTGCGRAAA 3286  
DB |||||  
DB 2930 AAGAGGTTCTCTTGGCATAAACAGGTTTCTGCCAATATGGATT--CCCAAGGTACCGCAAAA 2988  
QY 3287 TAGCCAGGYCATTAATATACASTAATTAAGGAACCTCAGAAAGCCCAATACCCATTTARTAA 3346  
DB |||||  
DB 2989 TAGCCAGACCGTTATATACGTTAATTAAGGAACCTCAGAAAGCCCAATACCCATTTAGTAA 3048  
QY 3347 GATGGAYAMCTGAAGYMRAGTGGCTTCCAGGCCCTTAAAGAAAGCCCTTAAACCCAAAGY 3406  
DB |||||  
DB 3049 GATGGACACCTGAGCAGAGCAGCTTTCATGCCCTTAAGAAAGGCCCT--AATGCNAGC 3106  
QY 3407 CCCAGTGTAAAGYTTGCGCAACRGGCCAAGACTTTTSTTVAITYARTCAC----- 3454  
DB |||||  
DB 3107 CCCAGTGTAAAGCTTGAAGGGGCAAGACTTTTCTCATATGTCACAAAAACCAAAACA 3166  
QY 3455 -----AGAAAAACAGAAAYAGCTCTGAGGATCTTTACACAGRTCCBAGGAGCA 3505  
DB |||||  
DB 3167 AACAAAGAAAAAATGATAGCTCTTAAGAGTCTTGCCTGCTGAGGACCA 3226  
QY 3506 GCTTCCAAACCYRTGGCRVACCTGAAGGAAAYTGATGTAGTGCAAAAGGGTTGRCYTC 3565  
DB |||||  
DB 3227 GTTTCGAACCCGTGGCATCTCTGA--TAAGGAATTTGATGTAGTGGCAAGGGTTCAGCCTC 3285  
QY 3566 ATTTGTTTAYGGTGTGTCAGTGTGAGTGTGAGTGTGTAAGAGGATTTAAATAATAC 3625  
DB |||||  
DB 3286 ATTTGTTTACGGGTAGTGTGTCAGTGTGAGTGTCTGAAGGATTTAAATAATAC 3345  
QY 3626 AGGPRAGAGATCTTACTGTGTGAGCATCTCATGAGTGAAYRGATACCTCATCTGTAAAG 3685  
DB |||||  
DB 3346 AGGGAAGAAATCTTACTGTGTGAGCATCTCATGATGTGAATGGCATACTCATCTGTAAAG 3405  
QY 3686 GAGACTTGTGCTGTGAGCAACACCTTAAATRTCAGGCTCTTATCTTGAARGGC 3745  
DB |||||  
DB 3406 GAGACTTGTGCTGTGAGCAACACCTTAAATRTCAGGCTCTTATCTTGAARGGC 3465  
QY 3746 CAGTGTCTGACTGTGCACTTGTGCAACTTTAAACCCAGYCNACTTTTCAGAGCAATG 3805  
DB |||||  
DB 3466 CAATCTCGCACTGTGCACTTGTGCACTTTAACTGAGACATTTCTTCCACAAATG 3525  
QY 3806 AAGAAAAATAPAAATTAATCTGTCAACAAATTAATTTCTCAAACTTATGCACTCGAGGG 3865  
DB |||||  
DB 3526 AAGAAAAATAPAAATTAATCTGTCAACAAATTAATTTGCTCAAACTTATGCACTCGAGGG 3585  
QY 3866 ACCTTGTAGAGTTCCTTGTGACTGTATCCYAGCTTCAACTTGTATATCTGATGGAAGTCTC 3925  
DB |||||  
DB 3586 ACCATTTAGAGTTCCTTGTGACTGTATCCAACTT--NACTTGTATCTGATGGAAGATCC 3644  
QY 3926 TTTGTAGAAAAAGGACTTCGAAAAAGYGGGGTATGAGTGGTTCAGTGAATAGTGAATYTT 3985  
DB |||||  
DB 3645 TTTGTAGAAAAAGGACTTCGAAAAAGTGGGTATGAGTAGTTCAGTGAATAGTGAATYTT 3704  
QY 3986 GAAAGTATCCCTCACTCCAGGAATAGTGTCTGAGTGTGAGAGAACTAATAGCCCTCACT 4045  
DB |||||  
DB 3705 GAAAGTATCCCTCACTCCAGGAATAGTGTCTGAGTGTGAGGCAAACT-----TCACT 3755  
QY 4046 KGGGCACTAGGAATTAGGAGAAAGGAAAAAGGAAAAATATATATACAGACTCTTARTATGCT 4105  
DB |||||



```
Db 3756 CGGCACTAGAAATTAGGAGAGGAAAGGGTAATATATATACAGACTCTAAGTATGCT 3815
Qy 4106 YACCTAGTCNTCCATGCGCCATGMRGCAATATGSRAGAAAGGGAATTCCTAACTCYGAG 4165
Db 3816 TACCTAGTCCTCCATGCGCCATGCAAGTAATATGGAGAGAAAGGGAATTCCTAACTCTGAG 3875
Qy 4166 RGAACACCTATCAACATCAGGAGCCATAGGARAATTATTAATGCGGWTACAGAAACCT 4225
Db 3876 GGAACACTATCAACATCAGGAGCCATAGGGAATTAATTAATGCGGWTACAGAAACCT 3935
Qy 4226 ARAGAGTGGMAGTCTTACACTGCGGGGTTCATCANAAAGGAAAGRAAGGGAATASAA 4285
Db 3936 AAAGAGTGGCAGTCTTACACTGCGGGGTTCATCAGAAAGGAAAGGGAATAGAA 3995
Qy 4286 GRGAATGCCAAGCAKATATTAAGCMAAAGAGCTCAAGGCGAGGACCTCCATTAGAA 4345
Db 3996 AGGAACCGCCAGCAGATATTAAGCCAAAGAGCTCAAGGCGAGGACCTCCATTAGAA 4055
Qy 4346 ATGCTTATTAACCTTCCCTTAGTATAGGGTAATCCCTTCGCGGAAACCAAGCCCCAGTAC 4405
Db 4056 ATGCTTATAGAGGACCCCTAGTATAGGGTAATCCCTTCGCGGAAACCAAGCCCCAGTAT 4115
Qy 4406 TCAGCAGGAGAAACAGAAATGGGAACCTCACGAGG---CAGTTTCTCCCTCGGACGG 4462
Db 4116 TCAGAGAGAAATAGAAATGGGAATCTCACGAGGACATAGTTTCTCCCTCAGGATGG 4175
Qy 4463 TTAGCCACTGAAGAGGAAATATCTTTGCTGCAACTATCCAATGGAAATTAATTTAAA 4522
Db 4176 CAAGCACCAGAAACAAAGGAAATACCTTTGCTGCACTAACCAATGGAATTAATTTAAA 4235
Qy 4523 ACCCTTCATCAACCTTTCACCTTAGGATCGATAGACCCCATCARATGGGCAATCTTTA 4582
Db 4236 ACCACTCACCAAACTTTTCACTTATGCAATGATAGCAACCCATCAGATGGGCAATTTGTTA 4295
Qy 4583 TTTACTGGACAGGCGCTTTTCAAACTATCAAGCARATATKTCAGGGCTCTGTAA 4636
Db 4296 TTTACTGGACAGGACTTTTCAAACTCTCAAGCAGATAGTTGGGCGCTGTAAA 4349

RESULT 10
US-10-450-763-4430
; Sequence 4430, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 4430
; LENGTH: 7466
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1378)..(2271)
; OTHER INFORMATION: 84% homologous to multiple sclerosis associated retrovirus
; OTHER INFORMATION: polyprotein, accession number AF009668, Smith-Waterman Score=1279.
US-10-450-763-4430

Query Match 44.2%; Score 3309.6; DB 10; Length 7466;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 3417; Conservative 98; Mismatches 17; Indels 38; Gaps 12;
Qy 1 CAACAATCGGGATATAAAACCCAGGCATTCGAGCTGGCAACAGCAGCGCCCTTTGGGTGC 60
```

```
Db 3924 CAATGATCGGGATATAAAACCCAGGCATTCGAGCTGGCAACAGCAGCGCCCTTTGGGTCC 3983
Qy 61 CTTCCCTTTCTGATGGGAG--CTGTTTTCAATGCTATTTCACTCTATTAATTAATTTGCAACTG 118
Db 3984 CTTCCCTTTCTGATGGGAGCTCTGTTTTCAATGCTATTTCACTCTATTAATTAATTTGCAACTG 4043
Qy 119 CACTCTTCTGGTCCATGTTTCTTAACGCTCGAGCTGAGCTTTTGTCTCACCCTCCACCACT 178
Db 4044 CACTCTTCTGGTCCATGTTTCTTAACGCTCGAGCTGAGCTTTTGTCTCACCCTCCACCACT 4103
Qy 179 GCTGTTTGGCCACACCGCANACCTGCGCTGACTCCCATCCCTCTGATCTCTGATCTCTGAGGGTG 238
Db 4104 GCTGTTTGGCCACACCGCAGACCTGCGCTGACTCCCATCCCTCTGATCTCTGAGGGTG 4163
Qy 239 TCCGCTGTGCTCTGATCCAGCGGARGCGCCATTTGCGCTCCCAATTTGGGCTAAAGGCTT 298
Db 4164 TCCGCTGTGCTCTGATCCAGCGGARGCGCCATTTGCGCTCCCAATTTGGGCTAAAGGCTT 4223
Qy 299 GCCATTGTNCTGCAACGCTTAAGTGTCTGGTTTGTCTTAATTTGAGCTGAACACTANTCA 358
Db 4224 GCCATTGTNCTGCAACGCTTAAGTGTCTGGTTTGTCTTAATTTGAGCTGAACACTANTCA 4283
Qy 359 CTGGGTTCCATGCTTCTCTGCTGACCCACGCGTCTTAAT-----AKAATATAA 409
Db 4284 CTGGGTTCCATGCTTCTCTGCTGACCCACGCGTCTTAATTTGAGCTGAACACTANTCA 4343
Qy 410 CACTTACCAATCGGCGCCAGATTTCCATTCCTTCGTAATCCGTGAGGSCAAACGAACTCCAGG 469
Db 4344 CACTTACCGCATGCGCCAGGTTCCGTTCTTGTAACTCCGTGAGGSCAA- GAACCCCGAG 4402
Qy 470 TCAGAGAATAACGAGCTTGCACCATCTTTGGAAGCGGCTGTACCTCTTTGGAAGTGGT 529
Db 4403 TCAGAGAATAACGAGCTTGCACCATCTTTGGAAGCGGCTGTACCTCTTTGGAAGTGGT 4462
Qy 530 TCACCAACCATCTTTGGGAGCTCTGTGAGCAAGGACCCCGCGGTACATTTTGGGACCAAMS 589
Db 4463 TCACCAACCATCTTTGGGAGCTCTGTGAGCAAGGACCCCGCGGTAACTTTTGGCAACCAAC 4522
Qy 590 RACGGACATCCMAAGTATGGGAAACGTTCCCGCAAGACAAACCGCCCTTAAGAGCTA 649
Db 4523 GACGGACATCCMAAGTATGGGAAACGTTCCCGCAAGACAAACCGCCCTTAAGAGCTA 4582
Qy 650 TTCTGGARAATTGGGAMCAATTTGACCTCAGACACTAAGAAAGAAACGACTTATATTTCT 709
Db 4583 TTCTGGARAATTGGGAMCAATTTGACCTCAGACACTAAGAAAGAAACGACTTATATTTCT 4642
Qy 710 TTCTGAGTGGCGCTGGCACTCTCTGAGGGAAGTATAAATTAATTAACCACTTTTACAGCTA 769
Db 4643 TTCTGAGTGGCGCTGGCACTCTCTGAGGGAAGTATAAATTAATTAACCACTTTTACAGCTA 4702
Qy 770 GACTCTTTTGTAGAAAGGCAAAATGGAGTGAAGTGCCTAAGTACAAACTTTCTTTTCA 829
Db 4703 GACTCTTTTGTAGAAAGGCAAAATGGAGTGAAGTGCCTAAGTACAAACTTTCTTTTCA 4762
Qy 830 TTAAGAGACAACTCAAAATTAATGTAATAAAGTGTATTTATGCCCTACAGAGACCTTCAG 889
Db 4763 TTAAGAGACAACTCAAAATTAATGTAATAAAGTGTATTTATGCCCTACAGAGACCTTCAG 4822
Qy 890 AGTCTACCTCCCTATCCAGCATTCCTTCCCAATTTTATTTTCCCAATTAAGAGACCTTCCTTC 949
Db 4823 AGTCTACCTCCCTATCCAGCATTCCTTCCCAATTTTATTTTCCCAATTAAGAGACCTTCCTTC 4882
Qy 950 RACCCAAATGGTCCAAAGGAGATAGACAAAGGTTAAACAGTGAACCAAGAGTGCCTAA 1009
Db 4883 AACCCAAATGGTCCAAAGGAGATAGACAAAGGTTAAACAGTGAACCAAGAGTGCCTAA 4942
Qy 1010 TATTTCCCAATTTATGACCTCTCCAGAGCTTAAAGCAATATAAAGAGAGAGTTCGCGCCAGCAGAG 1069
Db 4943 TATTTCCCAATTTATGACCTCTCCAGAGCTTAAAGCAATATAAAGAGAGAGTTCGCGCCAGCAGAG 5002
Qy 1070 TGCATGTGCTTTTCTTCCAGAGCTTAAAGCAATATAAAGAGAGAGTTCGCGCCAGCAGAGTTCCT 1129
```

Db 5003 TGCATGTGCTTTTCTCTCCAGACTTAAGCAATAAACAGACTTAGGTAAATCT 5062  
Qy 1130 CAGATAAACCCTGATGCTATATGRTGTTTTTCAAGGGTTAGGACAAATCTTTTGATCTGA 1189  
Db 5063 CAGATAAACCCTGATGCTATATGRTGTTTTTCAAGGGTTAGGACAAATCTTTTGATCTGA 5122  
Qy 1190 CATGAGAGATATATGTCATGCTTAATCAGACACTAAACCCCAAAATGAGAGAGTGC 1249  
Db 5123 CATGAGAGATATATATGTCATGCTTAATCAGACACTAAACCCCAAAATGAGAGAGTGC 5182  
Qy 1250 ACCATAAATGACGCTGAGGTTTGGCGATCTCTGGTATCTCAGTCAGTCAATGGATAN 1309  
Db 5183 ACCATAAATGACGCTGAGGTTTGGCGATCTCTGGTATCTCAGTCAGTCAATGGATAN 5241  
Qy 1310 GGATGACAAACAGAGAAAGANATGATTTCCCAACAGGCCAGARGCAGTTCCCAAGTCTA 1369  
Db 5242 GGATGACAAACAGAGAAAGAG-ATGATTTCCCAACAGGCCAGARGCAGTTCCCAAGTCTA 5300  
Qy 1370 SACCCTCATTTGGGGACACAGAAATCAGTAACATGGGAGATTGTCGTCAGACATTTGCT 1429  
Db 5301 GACCTCATTTGGGGACACAGAAATCAGTAACATGGGAGATTGTCGTCAGACATTTGCT 5360  
Qy 1430 AACTTGTGTGCTASAGGACTAAGGAAACTASGAGAAATCTAYGAATTACTCAATGA 1489  
Db 5361 AACTTGTGTGCTACAAGGACTAAGGAAATCTACGAGAAATCTACGAATTACTCAATGA 5420  
Qy 1490 TGTCCACCATTAACAAGGGAGAGGAGAAATTCCTACTGCTTCTTGAGAGACTAAGG 1549  
Db 5421 TGTCCACCATTAACAAGGGAGAGGAGAAATTCCTACTGCTTCTTGAGAGACTAAGG 5480  
Qy 1550 GAGGCAATCAGGAGGCTCTCTGTCACCTGACTCTTCTGAGGCCAACTAATCTTA 1609  
Db 5481 GAGGCAATCAGGAGGCTGCTCTGTCACCTGACTCTTCTGAGGCCAACTAATCTTA 5540  
Qy 1610 AAGCGTAAGTTTATCACTCAGTCAGTCAGACATTAAGAAAAAATCTCAAAAGTCTGCG 1669  
Db 5541 AAGCGTAAGTTTATCACTCAGTCAGTCAGACATTAAGAAAAAATCTCAAAAGTCTGCG 5600  
Qy 1670 TAGGCCCGGAGCAAACTTAGAAAACCTATTGAACCTTGCAACCTTTTAAATA 1729  
Db 5601 TAGGCCCGGAGCAAACTTAGAAAACCTATTGAACCTTGCAACCTTTTAAATA 5660  
Qy 1730 GAGATCAGAGGACGAGCGGACAGGACAAACGGGATTAAGAAAAAGGCCACCGCTTA 1789  
Db 5661 GAGATCAGAGGACGAGCGGACAGGACAAACGGGATTAAGAAAAAGGCCACCGCTTA 5720  
Qy 1790 GTCATGACCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAGGAGGAAAGCTGGGCAAT 1849  
Db 5721 GTCATGACCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAGGAGGAAAGCTGGGCAAT 5780  
Qy 1850 TGAATGCCCTAATAGGCTTGTCTCCAGTGGGCTCTAAGGACACTTTTAAAAAGATTGT 1909  
Db 5781 TGAATGCCCTAATAGGCTTGTCTCCAGTGGGCTCTAAGGACACTTTTAAAAAGATTGT 5840  
Qy 1910 CCAAGTAGAAGTAGAGCGGCTTCTGTCATGCGCTTATTTTCAAGGGAATCACTGGAAG 1969  
Db 5841 CCAAGTAGAAGTAGAGCGGCTTCTGTCATGCGCTTATTTTCAAGGGAATCACTGGAAG 5900  
Qy 1970 GCGCACTGCCAGGAGCAAAAGGCTTTTGTAGTCAGAGGCACTAAACAGATGATCCAG 2029  
Db 5901 GCGCACTGCCAGGAGCAAAAGGCTTTTGTAGTCAGAGGCACTAAACAGATGATCCAG 5960  
Qy 2030 CAGCAGGACTGAGGGTGTCTGGGGCAAGCGCATCCATGCCATCAACCTCTCAGAGCCC 2089  
Db 5961 CAGCAGGACTGAGGGTGTCTGGGGCAAGCGCATCCATGCCATCAACCTCTCAGAGCCC 6020  
Qy 2090 TGGGTATGCTTACCACTTACGAGGAGGAGGTTGTCCTCTGACACTGGTGGGCTTTC 2149  
Db 6021 TGGGTATGCTTACCACTTACGAGGAGGAGGTTGTCCTCTGACACTGGTGGGCTTTC 6080  
Qy 2150 TTAGTCTTACTCTTCTGTCGCGGACAACTGTCCTCAGATCTGTCATCTTCTGAGGGG 2209  
Db 6081 TTAGTCTTACTCTTCTGTCGCGGACAACTGTCCTCAGATCTGTCATCTTCTGAGGGG 6140

Qy 2210 TCCNTAAGACGGGCGAGTCAGTACTAGTACTTTTTTCCAGCCACTAAGTATGACTGGGAG 2269  
Db 6141 TCC-TAAGACGGGCGAGTCAGTACTAGTACTTTTTTCCAGCCACTAAGTATGAACTGGGAG 6199  
Qy 2270 CTTTATTTCTTTTACATGCTTTTCTAATTAATGCTTTGAAAGCCCACTACTCTTTGTTAGGA 2329  
Db 6200 CTTTATTTCTTTTACATGCTTTTCTAATTAATGCTTTGAAAGCCCACTACTCTTTGTTAGGA 6259  
Qy 2330 GAGACATTTCTAGCAAAAGCAGGGCCATTATACCTGAAATAGGAGAAACACCGG 2389  
Db 6260 GAGACATTTCTAGCAAAAGCAGGGCCATTATACCTGAAATAGGAGAAACACCGG 6319  
Qy 2390 TTTTGTGTCCTGCTTGTAGGAGGAGGATTAATCTGAGTCTGGCAACAGAGGACAA 2449  
Db 6320 TTTTGTGT-CCCTGCTTGTAGGAGGAGGATTAATCTGAGTCTGGCAACAGAGGACAA 6378  
Qy 2450 TATGACGAGCAAAAGGATGCGCTCTGCTTCAAGTTAAATCTAAGGATTTCCACTTCTTT 2509  
Db 6379 TATGACGAGCAAAAGGATGCGCTCTGCTTCAAGTTAAATCTAAGGATTTCCACTTCTTT 6438  
Qy 2510 TCCCTACCAAGGAGTACCCCTCAGACCCCAAGGCCCAACAGGATTTCCAAAAGATTGT 2569  
Db 6439 TCCCTACCAAGGAGTACCCCTCAGACCCCAAGGCCCAACAGGATTTCCAAAAGATTGT 6498  
Qy 2570 TAAAGACTTAAAGGCCCAAGGCTTAGTAAACCAATGCAATTAATCTCTGCAATTTCCGT 2629  
Db 6499 TAAAGACTTAAAGGCCCAAGGCTTAGTAAACCAATGCAATTAATCTCTGCAATTTCCGT 6558  
Qy 2630 AGTGATTTGAGGAGGACAGAAACCCAGTGCAGTGCAGGCTTAGTGCAAGATCTCAGG 2689  
Db 6559 AGTGATTTGAGGAGGACAGAAACCCAGTGCAGTGCAGGCTTAGTGCAAGATCTCAGG 6618  
Qy 2690 ATTATCAATGAGGCGTGTGCTTTTATACCCAGTCTGACTAGCCCTTATCTGCTGNY 2749  
Db 6619 ATTATCAATGAGGCGTGTGCTTTTATACCCAGTCTGACTAGCCCTTATCTGCTGCT 6678  
Qy 2750 TTCCCAATACAGAGGAGGAGGAGTGTGTACASTCTCTGAGCTTTMAGGATGCTTCTT 2809  
Db 6679 TTCCCAATACAGAGGAGGAGGAGTGTGTACASTCTCTGAGCTTTMAGGATGCTTCTT 6738  
Qy 2810 CTGCATCTCTGATCACTCTCAATCTTTTGTGCTTTGAGGATCTTTCAAAACCC 2869  
Db 6739 CTGCATCTCTGATCACTCTCAATCTTTGCTTTGAGGATCTTTCAAAACCC 6798  
Qy 2870 ARCATCTCAACTCACTGGAATTTTACCCAGGCTTCAAGGATGAGTCCCATCTATT 2929  
Db 6799 AACATCTCAACTCACTGGAATTTTACCCAGGCTTCAAGGATGAGTCCCATCTATT 6858  
Qy 2930 TGGCCAGGATTTAGCCCAAGGCTTGGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 2989  
Db 6859 TGGCCAGGATTTAGCCCAAGGCTTGGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 6916  
Qy 2990 GTAGTGTGATGATTTTACCTTTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 3049  
Db 6917 GTAGTGTGATGATTTTACCTTTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 6976  
Qy 3050 AGCRCTCTTMAATTTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3109  
Db 6977 AGCRCTCTTMAATTTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7036  
Qy 3110 CTGCTCAGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3169  
Db 7037 CTGCTCAGAGGATTTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 7092  
Qy 3170 TGAGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3229  
Db 7093 TGAGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7152  
Qy 3230 RGRRTTCTTGGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3289  
Db 7153 RGRRTTCTTGGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7211

```
QY 3290 CCAGGYCATTATACATACTAATTAAGGAAGAACTCAGAAAGCCAAATACCATTATTARTAGAT 3349
Db 7212 CCAGGTCATTAAATACACTAATTAAGGAAGAACTCAGAAAGCCAAATACCATTATTARTAGAT 7271
QY 3350 GGAYAMCTGAAGYMRAGTGGCTTTTCAGGCCCCCTTAAAGAGGCCCTTAAACCCCAAGYCCC 3409
Db 7272 GGACAACTGAAGTAGAAGTGGCTTTTCAGGCCC-----TAAACCCAGCCCC 7317
QY 3410 AGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTATATYRTCAAGAAAAAACAAGAA 3469
Db 7318 AGTGTTAAGYTTGCCAACRGGGCAAGACTTTTCTCATATGTACACAG-AAAAAACAAGAA 7376
QY 3470 YAGCTCTRGAGTCCCTTACACAGTCCRAGGAYGAGCTTGCAACCYRTGGCRYACCTGA 3529
Db 7377 TACCTCTAGAGTCCCTTACACAGATCCGAGGATGAGCTTGCAACCTGTGGCGTACCTGA 7436
QY 3530 STAAGGAAYTGATGTAGTGGCAAGGGTT 3559
Db 7437 CTAAGGAATGTAGTGGCAAGGGTT 7466

RESULT 11
US-11-121-086-79
; Sequence 79, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138 6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 79
; .LENGTH: 177175
; .TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-79

Query Match 41.5%; Score 3105.4; DB 15; Length 177175;
Best Local Similarity 86.7%; Pred. No. 0;
Matches 3580; Conservative 146; Mismatches 313; Indels 88; Gaps 29;

QY 606 GATGGGAACGTTCCTCCGCAAGCAAAAGAGCCCTTAAGACGTATTCTGGAATTTGGGA 665
Db 140106 GATGGGAATGTTCCTCCGCAAGCAAAAGAGCCCTTAAGATGTATTCTGGAGAAATGGGA 140165
QY 666 MCAATTTGACCTTCAGACACTAAGAAAGAAACGACTTATTCTCTGCAAGTCCGCTG 725
Db 140166 CCAATCTGACCATCAGACACTAAGAAAGAAAGACTTATTCTCTGCAAGTACCCTG 140225
QY 726 GC-----ACTCTGAGGGAAGTATAAATATAA 753
Db 140226 GCCAATATCTTCTTCAAGGGGCGAGAAACCTGGCTCTCTGAGGAAGTATAAATATAA 140285
QY 754 CACCATCTTACAGCTAGACATCTTTTGTAGAA---AAGCAAAATGAGGTGAAGTGCATA 810
Db 140286 CACCATCTTACAGCTAGACATCTTTTGTAGAAAGAAAGCAAAATGAGGTGAAGTGCATA 140345
QY 811 AGTACAACTTTCTTTTCAATTAAGACAACTCACAATTTATGTAAAAGTGTGATTATG 870
Db 140346 TGTACAACTTTCTTTTCAATTAAGACAACTCACAATTTATGTAAAAGTGTGATTATG 140405
QY 871 CCTTACAGGAAGCCTTCAGATCTTACCTCCCTATCCAGCA--TCGCCGACTCTTTCCTCC 928
Db 140406 CCTTACAGGAAGCCTTCAGATCTTACCTCCGACCCAGCAAGACCCCAACTCTTCTCC 140465
QY 929 AMYTAAATAGGACCCCTTCAACCCAAATGTGTCCAAAAGGAGATAGACAAAAGGGTAAA 988
Db 140466 AACTAATAGGACCCCTTCAACCCAAATGTGTCCAAAAGGAGATAGACAAAAGGGTAAA 140525
```

```
QY 989 CAGTGAACCAAGAGTGCCTAATATTCCTCAATTTATGACCCCTCCCAAGCAGTGGAGGAA 1048
Db 140526 CAATGAACCAAGAGTGCCTAATATTCCTCAATTTATGACCCCTCCCAAGCAGTGGAG 140581
QY 1049 GAGAATTCGGCCCGCCAGCAGTGCATGTGCTTTTTCCTCCAGACTTAAAGCAATAA 1108
Db 140582 GAGAATTCGGCCCGCCAGCAGTGCATGTGCTTTTTCCTCCAGACTTAAAGCAATAA 140641
QY 1109 AAAACAGACTTAGTGAATTTCTCAGATAAYCCTCATGCTATATTTGRTGTTTTCAAGGGT 1168
Db 140642 AAATAGACCTTAGTGAATTTCTCAGATAACCTCATGCTATATTTGRTGTTTTCAAGGGT 140701
QY 1169 TAGGACAATTTCTTGTATCTGACATGGAGAGATATATGTCTGCTTAATCAGACACTA 1228
Db 140702 TAGGACAATTTCTTGTATCTGACATGGAGAGATATA-ATGTTACTGCTTAATCAGACACTA 140760
QY 1229 ACCCCAAATAGAGAAAGTGCACCATTAACCTGAGCCTGAGRGTTTGGCGATCTCTGAT 1288
Db 140761 ACCCCAAATAGAGAAAGTGCACCATTAACCTGAGCCTGAGRGTTTGGCGATCTCTGAT 140820
QY 1289 CTCAGTCAGTCAATGATANGGATGACAAAGAGGAAGANAATGATTCCTCCACAGGC 1348
Db 140821 CTCAGTCAGTCAATG-ATAGGATGACAAAG-ATGAAAGAGAAATGATTCCTCCACAGGC 140877
QY 1349 CAGCARGCAGTTCCCACTCTASACCTCATTTGGGGACACAGAAATCAGTAACATGGGAGA 1408
Db 140878 CAGCARGCAGTTCCCACTCTGAGCCTCATTT-AGGACACAGAA---TCAGAACTTGGAGA 140933
QY 1409 TTGGTCTCAGACATTTGCTAACTTTGTGTCTASAGGACTAAGGAAAACCTAGAGAA 1468
Db 140934 TTGGTCCACACAGATTTGCTAACTTTGCTGTCTAGAGGACTAAGGAAAACCTAGAG-A 140992
QY 1469 AATCTAYGAATTAATCAATGATGTCCACCATACACAGGGGAGGAAGAAATCTCTACT 1528
Db 140993 AGCCCAATGAATTAATCAATGATGTCCCTATAACACA-GGGAAAGGAAGAAATCTCTACT 141051
QY 1529 GCCTTTCTGGAGAGACTAAGGAGGCAATTTAGGAAGCGCTGCTCTGTCACCTGACTCT 1588
Db 141052 GCCTTTCTGGAGAGACTAAGGAGGCAATTTAGGAAGCAATCTCTGTCACCTGACTCT 141111
QY 1589 TCTGAAGGCCAACTAATCTTAAAGCGTAAAGTTTATCAGTCAGTCAGTCAGACATT-AG 1647
Db 141112 ATTAAAGGCCAACTAATCTTAAAGGATAAGTTTATCAGTCAGTCAGTCAGAGATTAG 141171
QY 1648 AAAAAGCTTCAAAGTCTGCCCTAGGCCCGGAGCAAACTTAGAAAACCTATTGAACTTG 1707
Db 141172 AAAAAGCTTCAAAGTATGCTTTAGGCCCGGAGCAAACTTAGAAAACCTATTGAACTTG 141231
QY 1708 GCAACVTCGGTTTTTATAATAGATCAGGAGGAGCGCGGAGCAAGCAACCGGAT 1767
Db 141232 GCACCTCAGTTTTTATAATAGATCAGGAGGAGCGAGG-GGAAATGGGACAAATGGGAT 141290
QY 1768 -----TAAAAAAAAGGCCACCGCTTTAGTCAATGACCTCAGGCAAGTGGACTTTGGAG 1820
Db 141291 AAAAAAAAAGGCCAAAGTGACTGCTTTAGTCTGGCCCTCAGGCAAAATGGACTTTGGAG 141350
QY 1821 GCTCTGAAAAGGGAAAAGCTGGGCAAAATTTGAATGCCTAATAGGGCTTCTTCCAGTGGC 1880
Db 141351 GCTCTGAAAAGGGAAAAGCTGAGCAAAATTTGAATGCTTAAAGGGCTTCTTCTAGTGTG 141410
QY 1881 GTCTACAGGACACTTTAAAAAGATTTGTCGAAGTAGAAGTAAAGCGCCCTTCGTCCAT 1940
Db 141411 GTCTACAGGACACTTTAAAAAGATTTGTCGAAGTAGAAGTAAAGCGCCCTTCGTCCAT 141470
QY 1941 GCCCTTTATTTCAAGGGAATCACTGGAAGGCCCTGCCCCAGGGGACCAAGGCTCTTTTG 2000
Db 141471 GCCCTTTATTTCAAGGGAATCACTGGAAGGCCCTGCCCCAGGGGAGATGAAGGCTCTCTG 141530
QY 2001 AGTCAGAAAGCCAACTAACCCAGATGATTCAGCAGCAGGACTGAGGGTGCCTTGGGGCAAGCGC 2060
Db 141531 AGTCAGAAAGCCAACTAACCCAGATGATTCAGCAGCAGGACTGAGGATGCTCCAGGGCAAGCGC 141590
```

QY 2061 CATCCATGCGCATCAACCCCTCACAGAGCCCTGGGTATGCTTGACCATTTGAGGGCCAGGAAG 2120  
 Db 141591 CAGCCCATGCGCATCAACCCCTCACAGAGCCCTGGGTATGCTTGACCATTTGAGGGCCAGGAAG 141650  
 QY 2121 GT--TGTCCTCTGAGCACTGGTGGCGTCTTCTTAAGTCTTACTCTTCTGTCCTGGGACAA 2177  
 Db 141651 TTCACTGTCTCTGGACACTGGTATGGCTTCTCACTGTCTTCACTGTCTCTGTCCTGGGACAA 141710  
 QY 2178 TGTCTCTCAGATCTGCTCACTATTCTGAGGGGTCTNTAAGAGCGGCACTCACTAGATACT 2237  
 Db 141711 TGTCTCTCAGATCTGCTCACTATCC---GAGGGGTCTTAGGAGAGACATTTCTAGCAAGCAGCTAGATAC- 141766  
 QY 2238 TTTTCCAGCCCACTAAGTTATGAAGCTGGGAGCTTTATTTCTTTTCCACATGCTTTTCTAAT 2297  
 Db 141767 TTCTCCAGCCCACTGAGTTGTG-ACTGGGAACTTTACTCTTTTCCACATGCTTTTCTAAT 141825  
 QY 2298 TATGCTTTGAAAGCCCACTACCTTTGTTAGGGAGAGACATTTCTAGCAAAAGCAGGGGCCAT 2357  
 Db 141826 TATGCTGTAAAGCCCACTCCCTTGTAGGAGAGACATTTCTAGCAAAAGCAGTGGGCCAT 141885  
 QY 2358 TATACACTGAACATAGAGAAAGGAACACCGTTTGTGTGTCGCCCTGCTTGGAGGAAGAA 2417  
 Db 141886 TATACACTGAACATAGAGAAAGGAACACCACTTTGTGT-CCCCCTGCTTGGAGAAAGAA 141944  
 QY 2418 TTAATCCTGAAGTCTGGGCAACAGAGGACATATAGGAGCGACCAAGAAATGCCCCGTCT 2477  
 Db 141945 TTAATCCTGAAGTCTAAGCAACAGAGGACATATAGAGAG-CAAGAAATGCCCATCT 142003  
 QY 2478 GTTCAAGTTAAACTAAAGGATTCACCTTTCTTTTCCCTACCAAGGACAGTACCCCCCTAGA 2537  
 Db 142004 GTTCAAGTTAAACTAAAGGATTCGCGCTTTTCCCTACCAAGGACAGTACCCCCCTAGA 142063  
 QY 2538 CCCAAGGCCCAACAGGATTCCAAAGATTTGAAGGACTTAAAGGCCCAAGGCTTAGTA 2597  
 Db 142064 CCCAAGGCCCAACAGGACTCCAAAATACTGTAAAGGACTTAAAGGCCCAAGGCTTAGTA 142123  
 QY 2598 AAACCATGCAATACCTCCCTGCAAGTAACTCCGTAGTGGATTGAGGAGGACAGAAACCCAG 2657  
 Db 142124 AAACCATGCAAGTACGCCCTTGCATATCTCC-----AACTGTAGGAGTACAGAAACCCAA 142176  
 QY 2658 TGGACAGTGGAGGGTTAGTGGCAAGATCTCAGGATTTCAATGGAGGCCGCTTGTCTTTTGA 2717  
 Db 142177 CGGACAGTGA-GGTTAGTGGCAAACTCTCAGGATTTCAAT-GAGACTGTTGTTCTCTA 142234  
 QY 2718 TACCAGCTGTACTAGCCTTATATCTGTGMYTTCCAAATACCAAGAGGACAGAGTGG 2777  
 Db 142235 TACCAGCTGAACCTTAAACCTTATATCTGTCTTTCCAAATACCAAGAGGACAGAGTGG 142294  
 QY 2778 TTTTCACTCTGACCTTGAAGTGGCTTCTTCTGATCCCTGTATCCTCTGACTCTCAA 2837  
 Db 142295 TATATAGTCTGGAACCTTAAAGATGCTTTCTTGTGATCCCTGTATCCTGATTTCCCAA 142354  
 QY 2838 TTTCTTTTGTCTTTGAAGATCTTCAAACCCACATCTCAACTCACTGGACTTTTGA 2897  
 Db 142355 TTTCTGTTTGTCTTTGAAGATCTTCAAACCTCAACTCAACTCACTGGACTTTTGA 142414  
 QY 2898 CCCCAAGGGTTCAGGATGAGCCCACTATTTTGGCAGGCAATTAGCCCAAGACTTGAGY 2957  
 Db 142415 CGCCAAGGGTTCAGGATGAGCCCACTATTTTGGCAGGCAATTAGCCCAAGATTTGAGC 142474  
 QY 2958 CARTYMTACTCTGGACACTCTTGTCTTCTGTAGTGGATGATTTACTTTTGTGCGC 3017  
 Db 142475 CAAATCTCTACTCTGGACACTCTTGTCTTCTTGTATGTTGATGATTTACTTTTAGCCACC 142534  
 QY 3018 YRTTCAGAAACCTTTGGCCATCAAGCCACCAAGGCTCTTMAATTTCTTCCYACCTGT 3077  
 Db 142535 TGTTCAGAAACCTTTGGCCATCAAGCCACCAAGGCTCTTMAATTTCTTCCYACCTGT 142594  
 QY 3078 GGCTACAWGGTTTCCAAACASARGCTCACTCTGCTCAAGCAGGTTAAATCTTAGGR 3137  
 Db 142595 GGCTACAWGGTTTCCAAACCAAGGCTCAGCTCTGCTCAAGCAGGTTAAATCTTAGGG 142654  
 QY 3138 CTAARATTTCCAAAGGACCCAGGCCCTCAGTGGAGAAATATCCAGCCCTATCTGGCT 3197

Db 142655 CTAAAAATTATCCAAAGGCTTGGGGCCCTCAGTGAGGAACACATCCAGCCCTATCTGGCT 142714  
 QY 3198 TATCTCTCATCYCAAAAACCTTAAGCAACTAAGRRRTTCTTGGGCTTAAYAGGYTTCTGC 3257  
 Db 142715 TATCTCTCTTCCAAAACCTTAAGCAACTAAGGGTTTCTTGGGCTTAACAGGTTTCTAC 142774  
 QY 3258 CGAAWATGGATTTCCCAAGTGTGCGRAAATAGCCAGGYCATTTAWATACASTAAATTAAGA 3317  
 Db 142775 CGAATGTGGATT-CCAGGTACAGCAAAATAGGAGATCATTTATATACACTAATTAAGA 142833  
 QY 3318 AACTCAGAAAACCCCAATACCCATTTTARTAAGATGGAYAMCTGAAGYMRAGTGGCTTTCCA 3377  
 Db 142834 AACTCAGAAAACCCCAATACCCATTTTARTAAGATGGACACCTGAAGCAGAGCAGCTTTCCA 142893  
 QY 3378 GGCCCTTAAGAAAGGCTTTAAACCAAGYCCAGTGTGAAGYTTGGCCAAACGGGGCAAGAC 3437  
 Db 142894 GGCCC--TAAAGAGGGGCCCTAAACCAAGCCCGAGTGTAAAGCTTGGCCAA-GGGGCAAGAC 142950  
 QY 3438 TTTTSTTVATAYRTACAGAAAGAAAACAGRAAYAGCTCTGAGTCTTACACAGRTCCR 3497  
 Db 142951 TTTTCTTTATATGTCACAG-AAAAAACAGGAATAGCTCTAGAGGCCATCAACAAAGTCTG 143009  
 QY 3498 AGGAYGAGCTTGGCAACCCYRTGGCRYACCTGAATAAGGAAAYTGATGTAGTGGCAAGGG 3557  
 Db 143010 AGGACTAGCTTGGCAACCCGTTGGCATACTCTGATTAAGAAATTTGATGTAGTAAAGGA 143069  
 QY 3558 TTGRCYTCATTTGTTTAYGGGTAGTGGTGGAGTAGCAGTYKTAGTATCTGAAGCAGTTAA 3617  
 Db 143070 TTGGCTTCATTTGTTTACAGGTAGTGGTGGCAGTACAGTCTTAGTGTCTGAAGTGGTTAA 143129  
 QY 3618 AATAATCAGGAGAGAGATCTTACTGTGTGACATCTCATKAGTGAAYRGACATCTCAC 3677  
 Db 143130 AATGATACAGGAAGAGATCTTACTGTGTGACATCTCATGATGGAATGCCATCTCAC 143189  
 QY 3678 TGCTAAAGGACACTTGTGGCTGTGACAAACCGTGTACTTAAATRTCAGGCTCTTATCT 3737  
 Db 143190 TGCTAAAGGACACTTGTGGCTGTGACAAACCGTGTACTTAAATRTCAGGCTCTTATCT 143249  
 QY 3738 TGAARGCCAGTGTGCACTGTGCACTTGTGCAACTCTTAAACCCAGYCNATTTCTTCC 3797  
 Db 143250 TGAARGCCAGTGTGCACTGTGCAACTGTGTAACCTCTTAAACCCAGTCACTTTCTTCC 143309  
 QY 3798 AGCAATCAAGAAAGATARAAYATACTGTCAACAATAATTTCTCAACCTCTAGCCAC 3857  
 Db 143310 AGCAATCAAGAAAGATARAAYATACTGTCAACAATAATTTCTCAACCTCTAGCCAC 143369  
 QY 3858 TCGAGGGACCTTGTAGAGTTCCTTGTGACATCCYGAACCTTCAACTTGTATCTGATG 3917  
 Db 143370 TCGAGGGACCTTGTAGAGTTCCTTGTGACATCCCACTCTGATCTGATCTGATG 143428  
 QY 3918 GAAGTTCCTTTGTAGAAAAGGACTTCAAAAGYGGGGTATGCACTGTGTCAGTGTAGTAA 3977  
 Db 143429 GAAGTTCCTTTGTAGAAAAGGACTTCAAAAGYGGGGTATGCACTGTGTCAGTGTAGTAA 143488  
 QY 3978 GAATAYTTGAAGTAAATCCCTCACTCCAGGAACCTAGTGTCTGTCAGGAGAACTAATAG 4037  
 Db 143489 GAATAYTTGAAGTAAATCCCTCACTCCAGGAACCTAGTGTCTGTCAGGAGAACTAATAG 143548  
 QY 4038 CCYTCAYTTGGGCACTAGAAATTTAGGAGAAAGGAAAGGAAATATATATACAGACTCTR 4097  
 Db 143549 CCCTCACTCGGGTACTAGAAATCAGGAGAAAGGAAAGGAAATATATATACAGACTCTA 143608  
 QY 4098 ARTATGCTYACCTAGTCTNTCCATGCCATGMRGCAATATGSAAGAAAGGAAATCTTAA 4157  
 Db 143609 AGTGTGCTTACCTAGTCTCCATGCCATGCCATGCAATATGGAGAAAGGAAATCTTAA 143668  
 QY 4158 CTTTGAAGGAAACACTATCAACATCAGGAAGCCATTAGGARAATTTATTTAGGCGTAC 4217  
 Db 143669 CTTTCCGAGGAAACACTATCAACATCAGGAAGCCATTAGGAAAATTTATTTAGTGGTGTAC 143728  
 QY 4218 AGAAACCTTARAGAGGTGGAGTCTTACTGCTGCTGGGCTCATCANAAAGGAAAGGAAAGG 4277





Db	10784	GTCTACAAGGACACATTTTAAAAAAGATTGTCCAAGTAGAAAAAAGCTGCCCTTGTCCAT	10843
Qy	1941	GCCTCTTATTTCAAGGGATCACTGGAGGCCCACTGCCCAAGGGGACAAAGGCTCTTTTG	2000
Db	10844	GCCTCTTATGTCAAGGGAATCACTGGAGGCCCACTGCCCAAGGGATGAAGGCTCTCTG	10903
Qy	2001	AGTCAGAGGCCACTTAACAGATGATCCAGCAGCAGACTGAGGGTGCCTGGGGCAAGGC	2060
Db	10904	AGTCAGAGGCCACTTAACAGATGATCCAGCAGCAGACTGAGGATGCCAGGSCAAGGC	10963
Qy	2061	CATCCATGCCATCAACCTTCACAGAGCCCTGGGTATGCTTGACCAATTGAGGGCCAGGAG	2120
Db	10964	CAGCCCATGCCATCAACCTTCACAGAGCCCTGGGTATGCTTGACCAATTGAGGGCCAGGAG	11023
Qy	2121	GT---TGTCCTCGACACTGGTGGGTCTTCTTAGTCTTACTCTTCTGTGCCGACAAAC	2177
Db	11024	TTCACTGTCTCTTGGACACTGTATGGCCCTTCTCAGTCTTACTCTCTCTGTCTGTGACAAAC	11083
Qy	2178	TGTCCTCCAGATCTGTCACATATTCTGAGGGGTCCNTAAGACGGCAGTCACATAGATACT	2237
Db	11084	TGTCCTTTCAGATCTGTCACATATCC---GAGGGGTCTTAGGACAGCCAGTCACATAGATAC-	11139
Qy	2238	TTTCTCCAGCCACTAAGTTATGAACCTGGGAGCTTTATTTCTTTTTCACATGCTTTTCTAAT	2297
Db	11140	TTCTCCAGCCACTGAGTTGTG-ACTGGGGAACTTTACTCTTTTTCACATGCTTTTCTAAT	11198
Qy	2298	TATGCTTGAAGCCCACTACCTTGTTAGGGAGAGACATTTCTAGCAAAAGCAGGGGCCAT	2357
Db	11199	TATGGTGAAGCCCACTCCCTTGTTAGGGAGAGACATTTCTAGCAAAAGCAGTGGGCCAT	11258
Qy	2358	TATACACCTGAACATAGGAGAGAACCCGTTTGTGTCNCCCTGCTTGAAGGAGGAA	2417
Db	11259	TATACACCTGAACATAGGAGAGAACCCATTTGTGT-CCCCTGCTTGAGAAAGGAA	11317
Qy	2418	TTAATCCTGAAGTCTGGGCAACAGAGGACAAATATGAGCAGCGACAAAGAAATGCCGTCTCT	2477
Db	11318	TTAATCCTGAAGTCTAAGCAACAGAGGACAAATATAGACGAG-CAAGAATGCCATCCT	11376
Qy	2478	GTTCAAGTTAAACTAAGGATTCACATTCCTTTTCCCTACCAAGGAGTACCCTCTCAGA	2537
Db	11377	GTTCAAGTTAAACTAAGGATTCGCGCTCTTTTCCCTACCAAGGAGTACCCTCTTAGA	11436
Qy	2538	CCCAAGGCCCAACAGGATTCCAAAGATGTGTAGGACCTTAAAGCCCAAGGCTTAGTA	2597
Db	11437	CCCAAGGCCCAACAGGACCTCAAAGAACTGTTAAGGACCTTAAAGCCCAAGGCTTAGTA	11496
Qy	2598	AAACCATGCATAACTCCCTGCAGTAATTCCTGTAGTGGATTGAGGAGGCAACAGAAACCCAG	2657
Db	11497	AAACCATGCAGTAGCCCTGCAATACTCC-----AACTGTAGGAGTACAGAAACCCAA	11549
Qy	2658	TGGACAGTGGAGGTTAGTGCAAGATCTCAGGATTAATCAATGGAGCCCTTGTCTTTTGA	2717
Db	11550	CGGACAGTGGGA-GGTTAGTGCAAAATCTCAGGATTAATCAAT-GAGACTGTGTCTCTTA	11607
Qy	2718	TACCCAGCTGACCTAGCCCTTATATCTGTGMYTTCCTCCAAATACCCAGGAGCAGAGTGG	2777
Db	11608	TACCCAGCTGAACCTTAACCTTTATATCTGTCTTCCCAATACCAGAGGAGCAGAGTGG	11667
Qy	2778	TTTACASTCTCGACCTTMAAGATGCTCTTCTTCGATCCCTGTACATCTCACTCACTCAA	2837
Db	11668	TATATAGTCTTGAACCTTAAAGATGCTTCTTCTGATCCCTGTACATCTTGNATCCCAA	11727
Qy	2838	TTCTTGTGTTGCCCTTGAAGATCTTCAAACCCARCATCTCAACTCACTGGACTRTTTTA	2897
Db	11728	TTCTGTTTGGCTTTGAAGATCTTCAAACCTCAACGTCCTCAACTCACTGGACTGTTTTA	11787
Qy	2898	CCCCAGGCTTCAGGATAGYCCCAATCTATTTTGGCCAGGCATTTAGCCCAAGCTTTGAGY	2957
Db	11788	CGCCAAAGGTTACGGGATAGCCCCCATCTATTTTGGCCAGGCATTTAGCCCAAGATTTGAGC	11847
Qy	2958	CARTYMTACATCTGACACTCTTGTCTTTCRGTAAGTGGATGATTTACTTTTTRGCGCC	3017
Db	11848	CAATTCTCATACCTGGACACTCTTGTCTTGGTATGTGATGATTTACTTTTAGCCACC	11907
Qy	3018	YRTTCAGAAACCTTGTGCCATCAAGCCACCACAGCRCTCTTMAATTTTCTCGCYACCTGT	3077
Db	11908	TGTCAGAAACCTTGTGCCATCAAGCCACCACAGTGTCTTAAATTTCTTACCACCTGT	11967
Qy	3078	GGCTACAGGGTTTCCAAACSAARAGCTCARCTCTGCTCACAGCAGTTTAAATPACTTAGGR	3137
Db	11968	GGCTACAGGGTTTCCAAACCAAGGCTCAGCTCTGCTCACAGCAGTTTAAATPACTTAGGG	12027
Qy	3138	CTAAPATTATCCAAAGGACCARGCCCTCAGTGTAGGAAYATCCAGCTTACTTGGCT	3197
Db	12028	CTAAATTTATCCAAAGGCTTGAAGGCCCTCAGTGTAGGAACACATCCAGCTTACTTGGCT	12087
Qy	3198	TATCTCTATCYCAAAACCCCTAAAGCAACTAAGRBRRTTCTTGGCTTAAYAGGYTTCTGC	3257
Db	12088	TATCTCTTCCAAACCCCTAAAGCAACTAAGAGGGTTCTTGGCATACAGGTTTCTAC	12147
Qy	3258	CGAAWATGGATTCCCAGGTGTGGCRAAATAGCCAGGYCATTTAWATACASTAATTAAGGA	3317
Db	12148	CGAATGTGGATT-CCAGGTACAGCAAAATAGGCAGATCATTTATATACATTAATTAAGA	12206
Qy	3318	AACTCAGAAAGCCCAATACCCATTTARTAAGATGGAYAMCTGAAGYMRAGTGGCTTTCCA	3377
Db	12207	AACTCAGAAAGCCCAATACCCATTTAGTAAGATGGACACCTGAAGCAGAGCAGCTTTCCA	12266
Qy	3378	GGCCCTTAAGAGGSCCTTAAACCAAGVCCAGTGTAAAGVTTGCCAACRCGGCAAGAC	3437
Db	12267	GGCCC--TAAAGAGGGCCCTTAACCCAGCCCGTGTAAAGCTTGCCAA-GGGGCAAGAC	12323
Qy	3438	TTTTTSTTATATYRTCACAGAAAAAAACAGRAAYAGCTCTTRGGAGTCTTTACACAGTCCR	3497
Db	12324	TTTTTCTTATATGTCACAG-AAAAAACAGGAATAGCTCTAGGAGCCATCACAACAGTCTG	12382
Qy	3498	AGGAYAGCTTGCACCCVRTGGCRYACCTGASTAAGGAAATGATGTAGTGGCAAGGG	3557
Db	12383	AGGACTAGCTTGCACCCCGTGGCATACCTGAGTAAAGAAATGATGTAGTAGCAAGGA	12442
Qy	3558	TTGRCVTCATTTGTTTAYGGGTAGTGGTGGCAGTAGCAGTGTAGTATCTGTAAGCAGTTAA	3617
Db	12443	TTGGCTTTCATTTGTACAGGTAGTGGTGGCAGTAGCAGTCTTTAGTGTCTGAAGTGTTAA	12502
Qy	3618	AATAATACAGGAGAGATCTTACTGTGTGGACATCTCATGAGTGAAYRGATACACTCAC	3677
Db	12503	AATGATACAGGAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGCCATCTCAC	12562
Qy	3678	TGCTAAAGAGACTTGTGGCTGTGACAAACVGTGTTTACTTTAAATRTCAGGCTCTATTACT	3737
Db	12563	TGCTAAAGAGACTTGTGGCTGTGACAAACCGTTTACTTTAAATATCAGGCTCTATTACT	12622
Qy	3738	TGAAGGCCAGTGTGCRACCTGTGCACCTTGTGCAACTCTTAAACCCAGYCNCATTTCTTCC	3797
Db	12623	TGAAGGCCAGTGTGCAACTGTGCACCTGTGTACTCTTAAACCCAGTCAACATTTCTTCC	12682
Qy	3798	AGACAATGAAGAAAAAGATARAAYATAACTGTCAACAAATAATTTCTCAAAACCTATGCCAC	3857
Db	12683	AGACAATGAAGAAAAAGATARAAYATAACTGTCAACAAATAATTTCTCAAAACCTATGCCAC	12742
Qy	3858	TCGAGGGACCTTGTAGAGTTCCVTTGACTGATCCYGAACCTTCAACTTGTATCTGATGATG	3917
Db	12743	TCGAGGGACCTTGTAGAGTTCCVTTGACTGATCCYGAACCTTCAACTTGTATCTGATGATG	12801
Qy	3918	GAACTTCTTGTAGAAAAAGGACCTTGAAGAGYGGGTATGCAGTGTGTCAGTGAATAAG	3977
Db	12802	GAACTTCTTGTAGAAAAAGGACCTTGAAGAGYGGGTATGCAGTGTGTCAGTGAATAAG	12861
Qy	3978	GAAATTTGAAAGTAATCCCTCACTCCAGGAACCTAGTCTYAGCTTRGCAGAACTAATAG	4037
Db	12862	GAACTTGTAAAGTAATCCCTCACTCCAGGAACCTAGTCTCAGCTGGCAGAACTAATAG	12921
Qy	4038	CCYTCAATKGGGCACTAGAAATTTAGGAGAAAGGAAAGGGYAAATATATATATACAGACTCTR	4097
Db	12922	CCCTCACTCGGGTACTAGAAATCAGGAGAAAGGAAAGGGTAAATATATATATATACAGACTCTA	12981





Db 2710 GCAACCTCGGTTTTTTTATATAGAGATCAGGAGGAGCGGGAATGGACAAAGGGA 2769  
Qy 1767 TTAATAAAAGGCCACCGCTTTAGTCA TGA CCGCTCAGGCAAGTGAACCTTTGGAGGCTGTG 1826  
Db 2770 -TAAAAAAGGCCACCGCTTTAGTCA TGA CCGCTCAGGCAAGTGAACCTTTGGAGGCTGTG 2828  
Qy 1827 GAAAGGGGAAGAGCTGGGCAAAATGAAATGCCTAATAGGCTGTCTTCCAGTGGCGTCTAC 1886  
Db 2829 GAAAGGGGAATGCTTGAGCAAAATGGGTGCCTAA TAGGGCTTGTCTTCCAGTGCAGTCTAC 2888  
Qy 1887 AAGGACACTTTAAAAAGATTCTCAAGTGAAGTAAGCGGCCCTCTCGTCCATGCCCT 1946  
Db 2889 AAGGACATTTAAAAAGATTATCAAGTACAA TAGGCCACCCCTCGTCCATGCCCT 2948  
Qy 1947 TATTTCAAGGGGAATCACTGGAAGGCCACTGCCCCAGGGGCAAAAGGCTTTTGTAGTCAAG 2006  
Db 2949 TATATCAAGGGGAATCACTGGAAGGCCAGGGCCCCGGGGATGAAGGTCCTCTGAGTCAAG 3008  
Qy 2007 AAGCCACTAACAGATGATCCAGCAGCAGGACTGAGGCTGCGTGGGCAAGGCCATCCC 2066  
Db 3009 AAGCCACTAACAGATGATCCAGCAGCAGGACTGAGGCTGCGTGGGCAAGGCCATCCC 3068  
Qy 2067 ATGCCATCACTCAGAGGCCCTGGGTATGCTTGACCACTGAGGCCAGGAGGT--T 2123  
Db 3069 ATGCCATCACTCAGAGGCCCTGGGTATGCTTGACCTTTGAGGGCCAGGAGGTAACT 3128  
Qy 2124 GTCTCTCGACACTGGTGGGCTTTCTTAGTCTTACTTCTTGTCTCCCGCAACACTGTCT 2183  
Db 3129 GTCTCTCGACACTGGCAGCGCTTCTCAGTCTTACTGTCTCTGGACAACTGTCT 3188  
Qy 2184 CCAGATCTGTCACTATTCTGAGGGGTCNTAAGACGGGCAGTCACTAGATCTTTTCC 2243  
Db 3189 CCAGATCTGTCACTATCC---GAGGGGCTCTAGGACAGCAGTCACTAGATAC-TTCTCC 3244  
Qy 2244 CAGCCACTAAGTTATGAACCTGGGAGCTTTATTTCTTTCACATGCTTTTCTAATTATGCT 2303  
Db 3245 CAGGCACTAAGTTATG-ACTGGGAACTTTACTCATTTTCAATGCTTTTCTAATTATGCC 3303  
Qy 2304 TGAAGCCCCACTACCTTGTTAGGGAGACATTTCTAGCAAAAGCAGGGGCCATTATACA 2363  
Db 3304 TGAAGCCCCACTCCCTTGTAGTGAGAGAAATTTCTAGCAAAAGCAGGGGCCATTACACA 3363  
Qy 2364 CTTGAACATAGGAGAGGAACACCGTTTGTGTGTCCTGCTGTGAGGAGGAATTAATC 2423  
Db 3364 CTTGAACATAGGAGAGGAACACCGTTTGTGT-CCCCGTGTGAGGAGGAATTAATC 3422  
Qy 2424 CTGAAGTCTGGGCAACAGAGGACAAATATGGAGGACCAAGAAATGCCCGTCTGTTCAA 2483  
Db 3423 CTGAAGTCTGGGCAACAGAGGACAAATATGGAGGACCAAGAAATGCCCGTCTGTTCAA 3480  
Qy 2484 GTTAAACTAAGGATTTCCACTTCTTCCCTACCAAGGACGTAACCCCTCAGACCCAAAG 2543  
Db 3481 GTTAAACTAAGGATTTCCACTTCTTCCCTACCAAGGACGTA-CCCCCTAGACCCAAAG 3539  
Qy 2544 GCCCAACAAGGATTCOAAGATTGTTAAGGACTTAAAGCCCAAGCCCTAGTAAACCA 2603  
Db 3540 GCCCAACAAGGACTCCAAAGATTGTTAAGGACCGAAAGCCCAAGCCCTAGTAAACCA 3599  
Qy 2604 TGCATAACTCCCTGCAAGTAACTCCGTAGTGGATTGAGGAGGACAGAAACCCAGTGGACA 2663  
Db 3600 TGCATAAGCCCTTGCAATATTC-----AATTTTAGGATGACAGAAACCCAGTGGACA 3652  
Qy 2664 GTGGAGGGTGTAGTGCAAGATCTCAGGATTAATAGGAGCCGTTGCTCTTTTATACCCA 2723  
Db 3653 GTGGA-GGTTAGTGCAAGATCTCAGGATTAATCGAT-TAGGCTGTGTTCTCTATATCCA 3710  
Qy 2724 GCTGTACTAGCCCTTATCTGTGTTTCCCAATACACAGAGGAACAGAGTGGTTTACA 2783  
Db 3711 GCTGTACTAACCCCTTATACCTGCTTCCCAATATACAGAGGAACAGAGTGGTTTACA 3770  
Qy 2784 STCTCGACCTTMAAGATGCTTCTTCTGCATCCCTGTATCATCTGACTCTCAATTTCTG 2843  
Db 3771 GTCTGGACCTTAAGGATGCTTTTCTGCAATCCCTGTATCAATTTCTGACTCTCAATTTT 3830

Qy 2844 TTTGCTTTTGAAGATACCTTCAAAACCCARCATCTCAACTCACTGACGACTTTTTTACCCCAA 2903  
Db 3831 TTTGCTTTTGAAGATACCTTTCGAACCCAAATGCTTCAACTCACCTGACGACTTTTTTACCCCAA 3890  
Qy 2904 GGGTTACGGGATAGYCCCCCATCTATTTCGCCAGGCAATAGCCCAAGACTTGTAGYCAATYM 2963  
Db 3891 GGGTTACAGATAGCCCCCATCTATTTCGCCAGGCAATAGCCCAAGACTTGTAGCCAAATYC 3950  
Qy 2964 TCATACCTGGAACACTTTGTCTTCRGTAKGTGGATGATTTACTTTTTRGTCGCCVTTTCA 3023  
Db 3951 TCATACCTGGAACACTTTGTCTTCRGTAKGTGGATGATTTACTTTTGTAGCCCAATTTCA 4010  
Qy 3024 GAAACCTTGTGGCATCAAGCCCAACCAAGRCCTCTTMAATTTTCTCGCYACCTGTGGCTAC 3083  
Db 4011 GAAGCCTTGTGCCATCAAGCCCAACCAAGCACTCTTAAATTTTCTCGCTACCTGTGGCTAC 4070  
Qy 3084 AWGGTTTCCAAACSAARAGCTCARCTCTGCTCACAGCAGGTTAAATACTTAGRCCTAARA 3143  
Db 4071 AA-GATTCCAAACCAAGGCTCAGCTCTGCTCACAGAGTTTAATACTTAGGGCTAARA 4129  
Qy 3144 TTATCCAAAGGCAACARGGCCCTCAGTGAGGAAYATCCAGCCCTATCTGTGCTTATCTCT 3203  
Db 4130 TTATCCAAAGGTACCAGGGCCCTCAGTGGGGAACATATGACGCTATCTACTGGCTTGTCT 4189  
Qy 3204 CATCYCAAAACCTTAAGCAACTAAGRRRTTCTTGGCRTAAYAGGYTTCTGCCCAAWA 3263  
Db 4190 CATCCCAAAACCTT-AAGCAACTAAGAGGGTTCCTTGGCATTAACAGGTTTCTGCCAAATA 4248  
Qy 3264 TGGATTCCCAGGTGTGCRAAATAGCCAGGYCATTAATACASTAATAATTAAGCAAACTCA 3323  
Db 4249 TGGATT-CCAGGTATGGCAAAATAGCCAGATCATTTATATACACTTAATTAAGCAAACTCA 4307  
Qy 3324 GAAAGCCCAATACCCATTTTARTAAGATGGAYAMCTGAAGYMRAGTGGCTTTCCAGCCCC 3383  
Db 4308 GAAAGCCCAATACCCATTTTAGTAAGGTGGACACCTCAAGCACCAAGCAGCTTTCCAGGCTCT 4367  
Qy 3384 TAAAGAGGCTTAAACCAAGYCCAGTGTAAAGTTTGGCCAAACRGGGCAAGACTTTTTT 3443  
Db 4368 AAAGAA--GGCTCTAACCCAAAGTTCAGTGTTCACGCTTGGCCAAAGGGCAAGATTTTTCT 4425  
Qy 3444 TVATAYRTACAGAAAAAAGCAAGRAAYAGCTCTCGAGTCTTACACAGRTCCRAGGAY 3503  
Db 4426 TTAATGTACAGAAAAAAGCAAGAAATAGCTCTAGAGTCTTACACAGCTCCGAGGAT 4485  
Qy 3504 GAGCTTGCAACCCYRTGGCRYACCTGAATAAGGAAAYTGATGTAGTGGCAAGGGTTGRCY 3563  
Db 4486 GAACTTGCCAACTGTGGCATACCTGAGTAAGGAAATTGATGCAGTGGCAAGGGTTGGCC 4545  
Qy 3564 TCATTTGTTAYGGGTAGTGGCGAGTAGCAGTYKTAGTATCTGAGCAGTTAAAAATAAT 3623  
Db 4546 TCATTTGTTATGGGTAGTGGCGAGTAGCAGTCTTAGTATCTGAAGCAGTTAAAAATAAT 4605  
Qy 3624 ACAGGGRAGAGATCTTACTGTGGACATCTCATGAKGTGAAYRGCATACTCACTGCTAA 3683  
Db 4606 ACATGGAAGAGATCTTACTGTGGACATCTCATGATGTAAGCGGATACTCACTGCTAA 4665  
Qy 3684 AGGAGACTTGTGGCTGTCCAGCAACYGTTTACTTAAATRTCAGGCTCTAATTTCTGAARG 3743  
Db 4666 AGGAGACTTGTGGCTGTCCAGACGATCGTTTACTTAAACATCAGGCTCTAATTTCTGAAGG 4725  
Qy 3744 GCCAGTGTGBCACTGTGCACTGTGCAACTCTTAACCCAGYCNAGATTTCTTCCAGCAA 3803  
Db 4726 GCGTGTGTGCAACTGCGCGCCCTTTGCAACTCTTAAACCCAGCCACATTTCTTCCAGCAA 4785  
Qy 3804 TGAAGAAAAGATARAAYATAACTGTCAACAATTAATTTCTCAAACTATGCCACTCGAGG 3863  
Db 4786 TGAAGAAAAGATARAACATAACTGTCAACAAGTAATTTGTCAAACTA CCCCCTCGAGG 4845  
Qy 3864 GGACCTTTTAGAGTTTCCYTTGACTGATCCYGAACCTTCAACTTGTATCTGATGGAAGTT 3923  
Db 4846 GGACCTTTTAGAGTTTCCCTTGACTGATTCGACCTT- AACTTGTATATCTGATGGAAGTT 4904



QY	1588	TTCTGAAGGCCAACTAACTTTAAAGCGTAAGTTTATCATCTCAGTCAGCTGCGACATTAG	1647	QY	2662	CAGTGGAGGGTTAGTGCAGAGATCTCAGGATTTATCAATGGAGCGCGTGTCTCTTTTATACC	2721
Db	48728	TATTGAAGGCCAACTAACTTTAAAGAAAGTTTATCACTAGTCAGCTGCGAGATATTAG	48787	Db	49784	CAGTGER-GGTTAGTGCAGAGATCTCAGGATTTATCAAT-GAGGCCATTGTCCCTCTATACC	49841
QY	1648	-AAAAAACTTCAAAAGTCTGCGTAGGCGCCGAGCAAAACTTTAGAAACCTTATTGAACCTT	1706	QY	2722	CAGCTGTACCTAGCGCTTATATCTGTGMYTTCCCAAATACAGAGGAAGCAGAGTGTTTA	2781
Db	48788	AAAAAACTTCCAAAGTCGCTTAGGCCCGGAGCAAAAGTTTAGAAACCTTACTGAACCTT	48847	Db	49842	CAGCTGTACCTTAATCTTTATATCCGCTTCCCAAATCTAGAGGAAGCAAAAGTGTTTA	49901
QY	1707	GCACACCTGCGTGTATTAATAGAGATCAGAGAGAGCGAGCGGAACAGGACAAACGGGA	1766	QY	2782	CASCTCTGGACCTTMAAGATGCTTCTCTGCATCCCTGTGATCATCTCAGCTCTCAATTCT	2841
Db	48848	GGCAACCTCGGTGTATTAATAGAGATCAGAGAGAGCGGAGGAGCAAAATGGGA	48907	Db	49902	CAGTCTCTGGACCTTAAAGATGCTTTTCTGCACTCTTATACATGCTGACTCTCAATTCT	49961
QY	1767	T--TAAAAAAGGCCACGCTTTAGTCATGACCTCAGGCAAGTGGACTTTGGAGGCTC	1824	QY	2842	TGTTGCTTTGAAAGATACTTCAAAACCCARCATCTCAACTCAGCTGGACTTTTATACCC	2901
Db	48908	TAAAAAAGGCAAGCTGCTTTAGTCATGCGCTCAGGCAAGCGGACTTTGGAGGCTC	48967	Db	49962	TGTTGCTTTGAAAGATCTTGGAAACCAATCTCAACTCAGCTGGACTTTTATACCC	50021
QY	1825	TGAAAAAGGAAAGCTGGGCAAAATTGAATGCTTAATAGGGCTTGCTTCCAGTGGGTCT	1884	QY	2902	AAGGTTTCAGGGATAGYCCCATCTATTGTCAGGCAATAGCCCAAGACTTTGAGYCAT	2961
Db	48968	TGAAAAAGGAAAGCTGGGCAAAATAGATGCTTAATAGGGCTTGCTTCCAGTGGGTCT	49027	Db	50022	AAGGATTTCAGGGATAGCCCCCATCTATTGTCAGGCAATAGCCCAAGACTTTGAGCAGT	50081
QY	1885	ACAAGGACATTTAAAAAGATTGTCCAAGTAGAAGTAAGCCGCCCTTGTGCCATGCC	1944	QY	2962	YMTCACTCTGGACACTCTTGTCTTCTGTGTAGGTGATGATTTACTTTTGTGCGCYRTT	3021
Db	49028	-CAAGGACATTTAAAAAAGATTATCCAAATAGAAATAAGCCACTCCCTTGTCCATGCC	49086	Db	50082	TCTCATCTCTGGATATTCTTGTCTTGTGTATGCGGATGATTTACTTTAGCGCGCGT	50141
QY	1945	CTTATTTCAAGGAATCACTGAAGGCCACTGCCCGGAGGACCAAGGTCTTTGAGTC	2004	QY	3022	CAGAAACCTTGTGCCCATCAAGCCCAAGCRCTCTTMAATTTCTCGCYACTCTGGCT	3081
Db	49087	CTTATATCAAGGGAATCACTGTAAGGCCCACTGCCCGGAGGACGTAGGTCTCTGAGTC	49146	Db	50142	CAGAAACCTTGTGCCATCAAGCCCAAGGTGCTCTTAAATTTCTCGCCACTCTGGCT	50201
QY	2005	AGAAGCCACTAACACAGATGATCCAGCAGGACTGAGGGTGCTTGGGGCAAGCGCCATC	2064	QY	3082	ACAWGGTTTCCAAACSPARGCTCACTCTCTCTCAACAGCAG-----GTTAAA	3128
Db	49147	AGAAAGCCACTAACACAGATGATCCAGCAGGAGACTGAGGGTGCTTGGGGCAAGCGCCAGC	49206	Db	50202	ACAAAGTTTCCAAACCAAGGCTCAGCTCTGCTCAACAGCAGAGGGCTATTTATCCCTAAA	50261
QY	2065	CCATGCCATCACCTCACAGAGCCCTGGGTATGCTTGACCATGAGGGCCAGGAAGGT--	2122	QY	3129	TACTTAGGCTTAARATTTTCCAAAGGCACACGCGCCCTCAGTGAGGAAYRATCCAGCT	3188
Db	49207	CCATGCCATCACCTCACAGAGCCCTGGGTATGCTTGACCATGAGGGCCAGGAAGGTAA	49266	Db	50262	TACTTAGGGCTTAAATTTATCCAAAGGCACACGCGCCCTCAGTGAGGAATGTATCCAGCT	50321
QY	2123	-TGCTCTCGGACACTGGTGCGGTCTTCTAGCTTACTCTTCTGTCGCGGACACTGTC	2181	QY	3189	ATACTGGCTTATCTCATCYCAAAACCTAAAGCAACTAAGGRRRTTCTCT-----	3239
Db	49267	CTGTCTCTCGGACACTGGGTGCGGTCTTCTAGCTTACTCTTCTGTCGCGGACACTGTC	49326	Db	50322	ATACTGGCTTATCTTATCCCAAAACCTTAAACCAACTAAGNAGGTTCTTGGCATAATA	50381
QY	2182	CTCCAGATCTGTCACTATTCTGAGGGGTGCTNATAAGACGGGCAGTCACTAGATCTTTT	2241	QY	3240	-----GGCRTAAYAGVTTCTGCGGAATAATGATTTCCCGAGTTCGCCAGTAAATAGC	3290
Db	49327	CTCCATATCTGTCACTAT-----CCGAGCAGCCAGTCACTAGATAC-TTCT	49373	Db	50382	GGCATAACAGGCATAACAGGTTTCTGCTGAATATGGAAT-CCCAAGTACGGCAAAATAGC	50440
QY	2242	CCCAGCCATAAGTTATGAATCGGGAGCTTTATTTCTTTACATGCTTTTCTAATTATG	2301	QY	3291	CAGGYCATTAATACASTAATTAAGGAACCTCAGAAAGCCATACCCATTARTAGATG	3350
Db	49374	CCAGGCCATAAGTTGTG-ACTGGGGAACTTTACTGTTTTTCATGCTTTTCTAATTGTA	49432	Db	50441	CAGACCATATATACACTAATTAAGGAACCTCAGAAAGCCATATCCCAATTTAGTAGATG	50500
QY	2302	CTTGAAGCCCACTACTCTGTTAGGGAGAGACATCTAGCAAAAGCAGGCGCCATTATA	2361	QY	3351	GAYAMCTCAAGYMRAGTGGCTTTCCAGGCCCTTAAAGAGGCCCTTAAACCCCAAGYCCA	3410
Db	49433	CGTGAAGCCCACTCTCTTGTAGGGAGAGACATCTTAGCAAAAGCAGGCGCCATTATA	49492	Db	50501	GACACTGAAGCAGAGGAGCTTTCCAGGCCGTGAAGAACACCCCT--AAACCCAGGCCCA	50558
QY	2362	CACCTGAACATAGGAGAGGAACACCCGTTGTTGTCCTGCTTGAAGAGGAATTA	2421	QY	3411	GTGTTAAGYTTGCCAACRCGGGCAAGACTTTTSTTYATATYRTCAAGAAAAAACAGRAY	3470
Db	49493	CACCTGAACATAGGAGAGGAACACCCATTGTTGT-CCCTTCTGAGGAAGGAATTA	49551	Db	50559	GTGTTAAGCTTGCACGCGGGCAAGACTTTCTTTCTATGTATGTATGTATGTATGTATGT	50617
QY	2422	TCCCTGAAGTCTGGCAACAGAGGAACAATATGAGCAGAGCAAAAGAAATGCCCGTCTGTTT	2481	QY	3471	AGCTCTGGAGTCTTTACAGAGTCCAGGAGYAGCTTGCACACCTYRTGGGCRACCTGAS	3530
Db	49552	TCCTGAAGTCTGGCAACAGAGGAACAATACGGATGAG-CAAGAATGCCATCTGTTTC	49610	Db	50618	AGCTCTAGGAGTCTTTACACAGCTCCGAGGACCAAGCTTGAACCCATGCGCATACCTGAG	50677
QY	2482	AAAGTTAACTAAAGGATTCCTCTTCCCTTACCAAGGAGGACTACCCCTCAGACCCA	2541	QY	3531	TAAAGAAATGATGTAGTGGCAAAAGGTTGRCYTCATTTGTTTAYGGGTAGTGTGGCAGT	3590
Db	49611	AAAGTTAACTAAAGGATTCCTCTTCCCTTACCAAGGAGGACTACCCCTTAGACCGG	49670	Db	50678	TAAAGAAATGATGTAGTGGCAAAAGGTTGGCTCATTTGTTTACGGGTAGTGGGCGAGT	50737
QY	2542	AGGCCCAACAGGATTTCCAAAGATTGTTAAGGACTTAAAGCCCAAGGCTTAGTAAAC	2601	QY	3591	AGCAGTYKTATCTGAAGCAGTTTAAATATATACAGGGRAGAGATCTTACTGTGTGGAC	3650
Db	49671	AGGCCCAACAGGACTCCAAAGATTGTTAAGGACTTAAAGCCCAAGGCTTAGTAAAG	49730	Db	50738	AGCAGCTTAGTATCTGAAGCAGTTTAAATATATACAGGAAGAGATCTTACTGTGTGAAC	50797
QY	2602	CATGCATACTCCCTGCACTAAATTCCTGATGAGATTGAGGAGGCAAGAAACCCAGTGA	2661	QY	3651	ATCATCATGAGTGAAYRGCATCTACTGCTTAAAGGAGACTTGTGGCTGTAGACACACYG	3710
Db	49731	CATGCAGTAGCCCTGCACTACTCC-----AACTTTACGAGTACAGAAACCCACAGA	49783	Db	50798	CTCTCATGATGTGAACCCGACACTACTGCTTAAAGAAAGACTTGTGGCTGTGAGACACTG	50857
				QY	3711	TTTACTTAAATCTAGGCTCTATTACTTGAARGGCCAGTGTGTCGACTGTGCACTTGTGTC	3770

[illegible]

**RESULT 15**

US-10-461-862-12

US-10-461-862-12  
; Sequence 12, Application US/10461862

; Publication No. US20050090434A1

; GENERAL INFORMATION:

```

; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001800
; CURRENT APPLICATION NUMBER: US/10/461,862
; CURRENT FILING DATE: 2003-06-13
; NUMBER OF SEQ ID NOS: 184
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 156416
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-461-862-12

```

Query Match.	39.4%;	Score 2948;	DB 10;	Length 156416;
Best Local Similarity	85.0%;	Pred. No. 0;		
Matches 3523;	Conservative 144;	Mismatches 334;	Indels 144;	Gaps 32;
Qy	606	GATGGGAAACGTTCCCGCAAGACAAACAGCCCTTAAGACGTATTCTGGARAATTGGGA	665	
Db	55868	GATAGGAAATGTTCCCGCCCAAGGCAAAAATGCCCTTAAGATATATTCTGGAGAAATTGGGA	55927	
Qy	666	MCATTTTGACCTTCAGACACTAAGAAAGAAAGACGTATATTTCTCTGCAGTCGCCTG	725	
Db	55928	CCATTTTGACTCTCAGATCTTAAGAAAGAAATGACTTACATCTCTCGATGACACCTG	55987	
Qy	726	GC-----ACTCCTGAGGGAAGTATAAATTATAA	753	
Db	55988	GCCATGATGCTCTTCGAAGGGGAGAAACCTGGCCTCCTGAGGGAAGTATAAATTATAA	56047	
Qy	754	CACCATCTTACAGCTAGACYCTTTTGTAGAA---AAGGCAAAATGGAGTGCACATA	810	
Db	56048	CACCATCTTACAACTAGACCTCTTTGTAGAAAGAGGCAAAATGGAGTGAAGTGCACATA	56107	
Qy	811	AGTACAAACTTCTCTTTCATTAAAGAGACAACTCACAAATTATGTAAAAAGTCTGAATTATG	870	
Db	56108	TGTAACAACTTCTCTTTCATTAAAGAGACAACTTGCAATATGTAAAAAGTATGAATTATG	56167	
Qy	871	CCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCACAGCATCCCC--GACTCTCTCCCC	928	
Db	56168	CCCTACAGGAAGCCTTCAGAGTCTACCTCCCTAACCTGATGCTCCCTCTGACTCTCTCCCC	56227	
Qy	929	AMYTAAATAGGACCCGCC--TTCAACCCCAAATGGTTCAAAAGGAGATAGACAAAAGGGTAA	987	
Db	56228	AACATAAAGGACCCGCCCTTTCAACCCCAAACAGTCCAAAAGGACATAGACAAAAGGAGTAA	56287	
Qy	988	ACAGTGAACCAAGAGTGCCTAATATTCCTCCCAATATATGACCCCTCCCAAGCAGTGGGAGGA	1047	
Db	56288	ACAATGAACCAAGAGTGCCTAATATTCCTCGTTATATGCAACCTCC---AAGCGGTGGGA	56343	
Qy	1048	AGAGAAATCGGCCACGACGAGTGCATGTGCTTTTCTCCAGACTTAAGACCAATA	1107	
Db	56344	GAAGAAATCGGCCACGACGAGTGCATGCTTTTCTCTCTCACCTTGAACCAAAAT	56403	
Qy	1108	AAAAACAGACTTAGGTAAATTCAGATAAAYCCTGATGCTATATTTGRTGTTTTACAAGGG	1167	
Db	56404	AAAATAGACCTAGGTAAATTCAGATAAACCTTGATGCTATATTTGATTTTTACAAGGA	56463	
Qy	1168	TTAGGCAATCTTTGATCTGACATGGAGAGATATATATGTCATGCTTAATCAGACACT	1227	
Db	56464	TTAGGCAATCTTTGATCTGACATGGAGAGATATA-ATATTACTGCTTAATCATCAGACGT	56522	
Qy	1228	AAACCCAAATGAGAGAAGTGCCACCATAACTGACGCTGAGRGTTTCGCGATCTCTGTA	1287	
Db	56523	AACCTCAAATGAGAGAAGTGCTGCCATAAATCTGGAGCCCGAGAGTTTGGCAATCTCTGTA	56582	
Qy	1288	TCTCAGTCAGGTCAATGGATANGGATCAACACAGAAGGAAGANAAATGATTTCCCCACAGG	1347	
Db	56583	TCTCAGTCAGGTCAATG--ATAGGATGACAC--GGAGGAAGAGAACGAATCCCCACAGG	56639	
Qy	1348	CCAGCARGCAGTTCCTCAGTCTTASACCTTCATTTGGGGACACAGAAATCAGTAAACATGGGAG	1407	
Db	56640	GCAGCAGGCAAGTTCCTCAGTGTAGTCTCTCATTTGGGACACAGAAATCAGAGA---TGGAG	56695	







Db 58824 TAAGGAAATGATGATGAGTGGCAAGGGTTGGCCCTCATTTGTTACGGGTAGTGGCGGCAGT 58883  
Qy  
Db 3591 AGCAGTYKTATCTGTAAGCAGAGTTAAATAATACAGGGRAGAGATCTTACTGTGTGGAC 3650  
Db 58884 AGCAGTCTTAGTATCTGAGCAGTTAAATAATACAGGAGAGATCTTACTGTGTGAC 58943  
Qy 3651 ATCTCATGAKTGAAYRCATACCTACCTGTAAGAGGACCTTGGCTGTGACACAACYG 3710  
Db 58944 CTCTCATGATGTGAACCGCATCTACCTGTAAAGAGACCTGGTGGCTGTGACACAAC 59003  
Qy 3711 TTTACTTTAAATCTCAGGCTCTATTACTTTGAAGGCCAGTGTGTGACTGGCCTTGTGC 3770  
Db 59004 TTTGCTTAATATCAGGCTCTATTACTTTGAAGGCCAGTGTGTGACTGGCCTTGTGC 59063  
Qy 3771 AACTCTTAAACCCAGYCNCAATTTCTTCCAGACAATGAAGAAAGATARAAYATACTGTCA 3830  
Db 59064 AACTCTTAAACCCAGCAGCATTTCTTCCAGACAATGAAGAAAGATARAAYATACTGTCA 59123  
Qy 3831 ACAARTTAATTTCTCAAACTATGACCTCGAGGGGACCTTGTAGARGTTCCYTTGACTGA 3890  
Db 59124 ACAAGTAAATGCTCAAACTATGACCTCGAGGGGACCTTGTAGARGTTCCYTTGACTGA 59183  
Qy 3891 TCCYGACCTTCAACTTGTATACCTGATGGAAGTCTTGTAGAAAAGGACTTCGAAAAG 3950  
Db 59184 TCCCAACC-TCACTTGTATACCTGATGGAAGTCTTGTAGAAAAGGACTTCGAAAAG 59242  
Qy 3951 YGGGATATGACAGTGGTCAATGGAATATTTGAAGTAATCCCTCACTCCAGGAA 4010  
Db 59243 CAGAGTGTGTAGTGTGATGATGGAATATTTGAAGTAATCCCTCACTCCAGGAA 59302  
Qy 4011 CTAGTGTAGTGTGACAACTAATAGCCYTCAYTKGGGCACTAGAAATTAGGAGAA 4070  
Db 59303 CTAGTGTCTCAGCTGGCAGAACTAATAGCCCTCACTCAGGCACTAGAATTAGGAGAA 59362  
Qy 4071 AAAGGYAATATATATACAGACTCTATATGCTYACCTAGTCTCTCATGCCCATGMRG 4130  
Db 59363 AAAGGCAATATATATACAGACTCTAAGTATGCTTACCTAGTCTCTCATGCCCATGMRG 59422  
Qy 4131 CAATATGSRAGAAAGGAAATTCCTAATCTCYAGRGAACACCTATCAMACATCAGGAAG 4190  
Db 59423 CAATATGSGAAGAAAGGAAATTCCTAATCTCYAGRGAACACCTATCAMACATCAGGAAG 59482  
Qy 4191 CCATTAGGAATTAATTAATGTCMTGACAGAACTATAGAGAGTGGMACTTTACACTGCT 4250  
Db 59483 CCATTAGGAATTAATTAATGTCMTGACAGAACTATAGAGAGTGGMACTTTACACTGCT 59542  
Qy 4251 GGGGTCAATCANAAAGGAAGAAAGGAAATASAGRGAAYTGCCAAAGCAKATTTGAAG 4310  
Db 59543 G-----TAAGAAAGGACACAGAGAAATAAAGGGAACCCGCCAGTGGATATTGAAG 59591  
Qy 4311 CMAAAGAGCTGCAAGGACGACCCCTCCATTAGAAATGCTTATTAACCTTCCCTTAGTAT 4370  
Db 59592 CCGAAGAGCAACAGGCGGACCCCTCCATTAGAAATGCTTATTAAGAAACCCGCTAGTAT 59651  
Qy 4371 AGGATATCCCTTCGGGAAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGAA 4430  
Db 59652 GGGGTAATCCCTTCCAGAAACCAAGCCCGAGTACTCAGAAAGAAATAGAAATGGGAA 59711  
Qy 4431 CCTCAGGAGGAGTTTCTCC- CCTCGGACGGTTAGCCACTGAAAGAGGGAATACTT 4489  
Db 59712 CCTCATGAGGACGTAGTTCTCTCTCAGGATGGCTAGCCACCAAGAAAGGAAATACTT 59771  
Qy 4490 TTGCTGCACTATCCCAATGGAATTAATTAACCCCTTCAACCTTTCACCTTAGGC 4549  
Db 59772 TTGCTGCACTATCCCAATGGAATTAATTAACCCCTTCAACCTTTCACCTTAGGC 59819  
Qy 4550 ATCGATAGCAACCATCARATGGCCAAATCATTTATTACTGACCCAGGCTTTTCAAACT 4609  
Db 59820 ATTGATAGCAACCATCARATGGCCAAATCATTTATTACTGACCCAGGCTTTTCAAACT 59879  
Qy 4610 ATCAAGCARATKTCAGGCGCTGTGAATGTGCCARAAATAATCCCTGCCTTATCGC 4669

Db 59880 ATGAAGCAGATAGTCTAGAGGCTGTGAAGTGTGCAAAAAATAATCCCTGCACTTCAGGC 59939

Qy 4670 CAAGC 4674

Db 59940 CATGC 59944

Search completed: August 1, 2006, 19:39:10  
Job time : 5572 secs

**This Page Blank (uspto)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2006, 17:24:54 ; Search time 396 Seconds

(without alignments)  
9633.402 Million cell updates/sec

Title: US-10-717-580-11

Perfect score: 7489.6

Sequence: 1 caacaatcggtataaacc.....tattaatcttcgactgtr 7582

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1447667 seqs, 251571307 residues

Total number of hits satisfying chosen parameters: 2895334

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications NA.New.\*
- 1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
  - 2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
  - 3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
  - 4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
  - 5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
  - 6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
  - 7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
  - 8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	638.8	8.5	44063	US-10-539-228-718	Sequence 718, App
2	256.2	3.4	7323	US-10-539-228-718	Sequence 118, App
3	132.2	1.8	50000	US-10-528-659-2	Sequence 2, Appli
4	66.2	0.9	88493	US-10-539-228-703	Sequence 703, App
5	48.4	0.6	2579	US-11-218-305-4196	Sequence 4196, App
6	48.4	0.6	61791	US-10-539-228-645	Sequence 645, App
7	42.6	0.6	50	US-10-554-711-202	Sequence 202, App
8	42.2	0.6	6175	US-10-517-441-440	Sequence 440, App
9	42	0.6	3016	US-11-218-305-4195	Sequence 4195, App
10	41.6	0.6	50	US-10-554-711-192	Sequence 192, App
11	40.6	0.5	6175	US-10-517-441-714	Sequence 714, App
12	40.2	0.5	7975	US-11-256-173-12	Sequence 12, Appli
13	40.2	0.5	28605	US-10-471-571A-4495	Sequence 4495, App
14	39.2	0.5	1269	US-11-348-413-765	Sequence 765, App
15	37.6	0.5	2667	US-10-449-902-25901	Sequence 25901, A
16	37.6	0.5	155350	US-10-539-228-691	Sequence 691, App
17	37.2	0.5	913	US-11-174-307B-5089	Sequence 5089, App
18	36.8	0.5	1461	US-10-953-349-16839	Sequence 16839, A
19	36.8	0.5	1461	US-11-056-355B-59423	Sequence 59423, A
20	36.8	0.5	3943	US-10-449-902-16723	Sequence 16723, A
21	36.4	0.5	2040	US-11-348-413-6480	Sequence 6480, App
22	36.4	0.5	19634	US-10-517-441-688	Sequence 688, App
23	36	0.5	1153	US-11-218-305-7642	Sequence 7642, App
24	36	0.5	4944	US-11-218-305-1908	Sequence 1908, App
25	36	0.5	6001	US-10-517-441-772	Sequence 772, App

C	26	35.8	0.5	678	6	US-10-488-619-2734	Sequence 2734, Ap
	27	35.8	0.5	1604	6	US-10-449-902-7056	Sequence 7056, Ap
	28	35.6	0.5	220127	7	US-11-021-837-44	Sequence 44, Appl
	29	35.6	0.5	417	7	US-11-056-355B-75875	Sequence 75875, A
	30	35.6	0.5	100445	7	US-11-021-837-40	Sequence 40, Appl
	31	35.4	0.5	975	6	US-10-471-571A-3085	Sequence 3085, Ap
	32	35.4	0.5	1217	6	US-10-449-902-23455	Sequence 23455, A
	33	35.2	0.5	2391	7	US-11-348-413-6522	Sequence 6522, Ap
	34	35.2	0.5	5023	6	US-11-218-305-19413	Sequence 19413, A
	35	35.2	0.5	6680	6	US-10-517-441-574	Sequence 574, App
	36	35.2	0.5	7309	6	US-10-517-441-333	Sequence 333, App
	37	35.2	0.5	7309	6	US-10-517-441-607	Sequence 607, App
	38	35	0.5	2307	7	US-11-056-355B-88420	Sequence 88420, A
	39	35	0.5	2307	7	US-11-056-355B-92176	Sequence 92176, A
	40	35	0.5	2308	7	US-11-056-355B-28176	Sequence 28176, A
	41	35	0.5	2308	7	US-11-056-355B-31766	Sequence 31766, A
	42	35	0.5	2986	6	US-10-524-979-1	Sequence 1, Appli
	43	34.8	0.5	2245	7	US-11-056-355B-12956	Sequence 12956, A
	44	34.8	0.5	2614	7	US-11-218-305-7623	Sequence 7623, Ap
	45	34.8	0.5	19634	6	US-10-517-441-414	Sequence 414, App

#### ALIGNMENTS

##### RESULT 1

US-10-539-228-718

; Sequence 718, Application US/10539228

; Publication No. US20060154250A1

; GENERAL INFORMATION:

; APPLICANT: David W. Morris

; APPLICANT: Marc S. Malandro

; TITLE OF INVENTION: Novel Compositions and Methods in Cancer

; FILE REFERENCE: CHIR0052-101 (PP023370.0003)

; CURRENT APPLICATION NUMBER: US/10/539,228

; CURRENT FILING DATE: 2005-06-17

; PRIOR APPLICATION NUMBER: US 10/322,281

; PRIOR FILING DATE: 2002-12-18

; NUMBER OF SEQ ID NOS: 866

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 718

; LENGTH: 44063

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1) - (44063)

; OTHER INFORMATION: n = A,T,C or G

US-10-539-228-718

Query Match 8.5%; Score 638.8; DB 6; Length 44063;

Best Local Similarity 85.3%; Pred No 1e-185;

Matches 732; Conservative 41; Mismatches 65; Indels 20; Gaps 6;

QY 3690 CTTGGCTGTGTCAGAACACGTTTCTTAAATRTCAGGCTCTTACTTCTTAAAGGCGCAGT 3749

DB 1 CTTGGCTGTGTCAGAACACGTTTCTTAAATRTCAGGCTCTTACTTCTTAAAGGCGCAGT 60

QY 3750 GCTGCACTGTGCACTTGTGCACTCTTAAACCAGCAGCAGTCTTCCAGACAGATGAAGA 3809

DB 61 GTTGGCAATGCACTTGTGCACTCTTAAACCAGCAGCAGTCTTCCAGACAGATGAAGA 110

QY 3810 AAAGATARAAYATACTGTCAACAARTATTTCTCAACCTATGCATCCAGGGGACCT 3869

DB 111 AAAGACAGATACAATGTCAATAGGTGATTCCTCAACCTATGCATCCAGGGGACCT 170

QY 3870 TTTAGAGTTCCCTTGTGATCCGACCTTCAACTTGTATCTGATGAAGTTCCTTTG 3929

DB 171 TCTAGAGTTCCCTTGTGATCCGACCTTCAACTTGTATCTGATGAAGTTCCTTTG 229

QY 3930 TAGAAAAAGGACTTCGAAAAGGGGATGAGTGGTGTAGTGAATATGGAATATTTGAAA 3989

DB 230 TAGAAAAAGGACTTCGAAAAGGGGATGAGTGGTGTAGTGAATATGGAATATTTGAAA 289

```
QY 3990 GTAATCCCTCACTCAGGAAGTAGTCTYAGCTTGCAGAACTAATAGCCYTCATKGGG 4049
Db 290 GTAATCTCTCACTCAGGAAGTAGTCTCAGCTGGCAGAACTAATAGCCCTCACTCGGA 349
QY 4050 CACTAGAATTAGGAGAGAAAGGAAATATATATACAGACTCTTARTATGCTYACC 4109
Db 350 CACCAGAAATTAGGAGAGAAAGGAAATATATATACAGACTCTTARTATGCTYACC 409
QY 4110 TAGTCNTCCATGCCATGMRGCAATATGSARAGAAAGGAATTCCTAACTTCYAGAGAA 4169
Db 410 TAGTCCTCCATGCCATGAGCAATATGAGAGAAAGGAAATTCCTAACTTCYAGAGAA 469
QY 4170 CACTATCMACATCGGAAGCATTAGGARATTAATATGCGWGTACAGAAACCTARAG 4229
Db 470 CACTATCAAAATCAGGAAGCATTAGGAGATTAATCTTGGCTCTACAGAAACCTAAG 529
QY 4230 AGGTGMACTCTTACACTCYGGGTTCATCANAAAGGAAGRAAGGAAATASAGRA 4289
Db 530 AGGTGCGAGTCTTACACTCCAGGTCATCA-----GAAAGTAAAGGGAATAGAGGA 584
QY 4290 AYTGCCAACAKATATTGAAGCMAAAGAGCTGCAAGGCGAGGACCTCCATTAGAAATGC 4349
Db 585 ACTGCCAAGCGGATATTGAAGCC-AAAGATCCGCAAGGCGAGGACCTCCATTAGAAATGC 643
QY 4350 TTATTAACTCCCTTAGTATAGGTAATCCCTTCGGGAAACCAAGCCCGAGTACTCG 4409
Db 644 TTATAGAAGGATCCCTTAGTATAGGTAATCCCTTCGGGAAACCAAGCCCGAGTACTCG 703
QY 4410 CAGGAGAACAGAAAGGGAACCTCAGGAG--CAGTTTTCTCCCTCGGACGGCTAG 4466
Db 704 CAGAGAAATAGAACGGGGAACCTCAGGAGGACATAGTTCTCCCTCAGGATGCGTAG 763
QY 4467 CCACGAAAGAGGAAATACATTTTGGCTGCAACTATCCAAATGGAATTACTTAAACCC 4526
Db 764 CCACGAAAGAGGAAATACATTTTGGCTGCGGCTAATCAATAGAAATTAATGTAAGGCC 823
QY 4527 TTCATCAAACTTCACT 4544
Db 824 TTCACCAAACTTCACT 841
```

## RESULT 2

```
US-10-539-228-118/c
; Sequence 118, Application US/10539228
; Publication No. US20060154250A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: CHIR0052-101 (PP023370.0003)
; CURRENT APPLICATION NUMBER: US/10/539,228
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: US 10/322,281
; PRIOR FILING DATE: 2002-12-18
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 118
; LENGTH: 73723
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(73723)
; OTHER INFORMATION: n = A,T,C or G
US-10-539-228-118
```

```
Query Match 3.4%; Score 256.2; DB 6; Length 73723;
Best Local Similarity 45.4%; Pred. No. 6.1e-67;
Matches 1285; Conservative 83; Mismatches 1384; Indels 80; Gaps 17;
QY 2650 AAACCCAGTGGCAGTGGAGGGTTAGTGCAGAGATCTCAGAGTTATCAATGGAGCGGTG 2709
```

```
Db 58130 AABACCAGACAAGTCTTACAAAGTTAGTTCAGGATCTGGCCCTTATGAA-CCAAAATATTTT 58072
QY 2710 TCCTTTTATACCCAGCTGTACTAGCCCTTATATCTGTGMYTTCCAAATACAGAGGAG 2769
Db 58071 TACTTATCCACCCGCTGGTGCACCAACCATATATCTCTTTGTCTCAATACCTTCCCTCCA 58012
QY 2770 CAGAGTGGTTTACASTCTCTGACCTTMAAGGATCCCTTCTGTGCATCCTGTACATCTGT 2829
Db 58011 CAATCACTGTTCGCTTTTGTGATCTTAAAGATGCTTTTTTCACTATCCCTGCAACCT 57952
QY 2830 ACTCTCAATCTTGTGTGCTTTGAAGATACCTT-----CAAACCCARCATCTCAAC 2880
Db 57951 TATCCAGTCTCTCTTTGCTTTTACCTGGACTGACCCCTGACACCCATCAGTCCAGCAGT 57892
QY 2881 TCACCTGACATRTTTTACCCAAAGGTTACAGGATAGTCCCATCTATTTGGCCAGGAT 2940
Db 57891 TTACCTGGCTGTGTGCGCAAGCTTTCAGGAGCAGCCCTCATTTCTCAGCCAAGCTC 57832
QY 2941 TAGCCCAAGACTTGAGYCATYMTATACATCTGGACACTCTTGTCTTCRGTAKGTGGATG 3000
Db 57831 TTTCTCATGATTTACTTTCTTTCCACCCCTCCACTTCTTCACTTATCAATATATGATG 57772
QY 3001 ATTTACTTTTTCGCTGCTTTTTCAGAAACCTTTGTCATCAAGCCCAAGGCTCTTMA 3060
Db 57771 ACCTTCTACTTTGTAGTCCCTCTTTGAGTCTTCTCAACAGACACCCCTCTGCTCTTC 57712
QY 3061 ATTTCTCGCVACTGTGTGCTTACAWGTTTCCAAACARARGCTCACTCTGTCTCAGC 3120
Db 57651 CCGTTTACTGTCTTGGCATATTTCTTATATAAAACACACACTGTCTCTCCCTGCCATGTG 57592
QY 3181 TCCAGCTTATCTGTGCTTATCTCATCYCAAAACCCCTAA-AGCACTAAGRRRTTCTTT 3239
Db 57591 TCCAACTGTCTCTCAAAACCCCAACACTTCTCAAAACCAACACTCTCTTCTTCTCTA 57532
QY 3240 GGCRTAAYAGYTTCTGCCGAATATGATTTCCCGAGTTGTCGCAAAATAGCCAGGTCAT 3299
Db 57531 GGCATGTTGGATATCTTTCACCTTTGGATACCTG-GTTTGTGCATCTCTAAACAACCCATT 57473
QY 3300 AATACASATTAATAGGAACCTCAGAAAGCCAATACCCATTTATTAAGATGGAYAMCTGA 3359
Db 57472 ATGTAAATTCAAAAAGGAACCTAGCTGACCCCAT-----AGATCTTAATCT 57423
QY 3360 AGYMAAGTGGCTTTTCAGGCGCCCTTAAAGAGGCTTAAACCAAGYCCAGCTGTTAAGY 3419
Db 57422 TTCCCACTCTCTTCTTCCATTTCTTGAAGACAGATTTAGAGACTGCTCCACACTAGCTC 57363
QY 3420 TTGCCAACRGGGCAAGACTTTTSTTATATATATCAAGAAAAAAGCAAGAA-AGCTCTRG 3478
Db 57362 TCCCTGACTCATCCCAACCCCTTTGGTTACACACAGCTGAAGTGCAGGCTGTGCAGTTG 57303
QY 3479 GAGTCTTTACACAGTCCBAGGAYGAGTTGCCAACCTTGCRTACCTGASTAGGAAA 3538
Db 57302 GAAATCTTACACAGTACCGGAGCCGCGC-----CCTGTAGACCTTTTGTCAAACAAC 57249
QY 3539 VTGATGTAGTGGCAAGGCTTGRCTCATTTGTTTAYGGGTAGTGTGGCAGTAGCAGTYK 3598
Db 57248 TTGACCTTACTGTTTGTAGACTGGCCATGATGTCTCCATGACGCGGTGCCACTGCGCTAA 57189
QY 3599 TAGTATCTGAAGCAGTTTAAATATATACAGGAGAGATCTTACTGTGTGGACATCTCATG 3658
Db 57188 TACTTTTAGAGGCCCTTAAATATCAACACTATGTCTCACTGTCTACAGCTCTCGTA 57129
QY 3659 AKGTGAAYRGCATACTCACTGCTAAAGGAGACTG---TGGCTGTACAGACAACGTTTAC 3715
Db 57128 ACTTCCAAAATCTATTTCTTCTTCCACTGTGATGATATCTTCTGTCTCCCGAGCTCC 57069
QY 3716 TTAATATCTCAGCTCTATTTACTTGAAGGCCAGTGTGCTGCACTGTGCACTTGTGCAACTC 3775
Db 57068 TTCAGTTATCTCACTCTTTGTTGAGACTCCCACTTACCATTGTTCTCTGCGCCAGACT 57009
```

Qy	3776	TTAACCAGYCNCAITTTCTCCAGACATGAAGAAAGATARAATATCTGTCAACAAR	3835
Db	57008	TCATCCAGCCTCCCACAITATTTCCGGATACCACCTGCACCCCACTGATCTCTC	56949
Qy	3836	TAATTTCTCAAACTATGCCACTCGAGGGGACCTTGTAGARGTTCCYTTGACTGATCGVG	3895
Db	56948	TGATCCACCTGCACATACACCCCATTTCCCATGTTCTCTTCTCTGTTCTCAACCTG	56889
Qy	3896	ACCTTCAACTTGTATCTATGATGAAGTTCTTGTGTAGAAAAAGAGCTTGGAAAAGYGGGG	3955
Db	56888	ATC-ACATTTGGTTTACTGTATGCGACGCTCCATCAGGCCTTAATCGCCACTCACCAACAAG	56830
Qy	3956	TATGCACTGCTCAGTGATGAATGGA-----ATAVTTGAAAGTAATCCCTCTCACTCCAGGA	4010
Db	56829	GCAGGCTATGCTATATATCTTCCACATCTATCATTTGAGGCTACCGTCTGCCCTTACTCCA	56770
Qy	4011	CTAGTGTCTYAGCTFRGCAGAACTAATAGCCYTCAYTKGGGCACTAGAATTTAGSAGAGRAA	4070
Db	56769	CTACCTCTCAACAAGCGCAACTCATTTGCCTTAACTCGGGCCCTCACTCTTGTCAAAGGAT	56710
Qy	4071	AAAGGYYAATATATATACAGACTCTPABTATGCTYACCTAGCTYNTCATGCCCATGMRG	4130
Db	56709	TACATGTCAATATTTATGCTGACTCAAAATATGCCCTTCCACATCTCTGCACCACTATGCTG	56650
Qy	4131	CAATATGSARAGAAAGGGAAATTCCTAACTCTYAGRGAAACCTATCAMACATCAGGAAG	4190
Db	56649	TTATATAGGCTGAAAGAGGTTTCTCACTACACAGAGTCTCTCACTCATTAATGTCTTTT	56590
Qy	4191	CCATTAGGARATTTATYATGGCWGTAACAGAAACCTARAGAGGTGGMAGTCTTTACACTGCY	4250
Db	56589	TAATAAAAACTCTCTCAAGGCGCTTACTTCCAAAAGAAAGCTGGAGTCAATCACTGCA	56530
Qy	4251	GGGTCATCANAAAGGAAGRAAGGGAATASAGRGAATYGCCAAGCAKATATTGAAG	4310
Db	56529	AAGGCCATCAAAAGGCATCAGATCCCATTTGCCCTCGGCAACGCTTATGCTGATAGGTAG	56470
Qy	4311	CMAAAAGAGCTCAAGCGCAGGACCTCCATTAGAAATGCTATTAAACTTCCTCTTAGPAT	4370
Db	56469	CTAAGNAGCAGTAG-----TGTTCCAACTTCTGTCTCTCTACGGCC	56428
Qy	4371	AGGGTAATCCCTTCCGGGAAACCAAGCCCGAGTACTCAGCAGGAGGAAACAGAAATGGGAA	4430
Db	56427	AGTTTTTCTCTCTCATCGGTCACTTCCACCTACTCCCCCACTGAACTC-----	56377
Qy	4431	CCTCAGAGCGCACTTTTCTCCCTCGGAGGGTTAGCCACTGAAGAAGGGAATACTTTT	4490
Db	56376	CCAACCTGTCAATCTCTTTCCACATAGGCAAAATGGTTCTTAGACCAAGGAAGTATCTCC	56317
Qy	4491	TGCGTCAACTATCCAAATGMAATTACTTAAACCTTCATCAACACCTTTCACATTAGCA	4550
Db	56316	TTCCAGGCTCAGAGGCCCATTTCTATGTGTCTCATTTTTCATAACCTCTTCATGTAGTT	56257
Qy	4551	TCGATAGCACCCATCARATGGCCAAATCAATTTTACTGGACCGAGGCCCTTTTCAAAACTA	4610
Db	56256	ACAAGCGCTAGCCCGCTCTTAGAATCTCTAATTTCTTTTCCATCATGGAATCTGTCC	56197
Qy	4611	TCAAGCARATAKTCAGGGCCTGTGAATGTGTGCCARARAATAATCCCTCGCTTATCGCC	4670
Db	56196	TCAAGGAATCACTTCTCAGTGTTCATCTGCTATTCTACTACCCCTCAGGGATTTTCA	56137
Qy	4671	AAGCTCTCT-----TCAGGARAAACAARAACAGGCCATTTACCCTGARAARACTGGC-AA	4723
Db	56136	GGCCCCCTCCCTTCCCTACATCAAGCTCAGGACTTGGCCCTTCCCAAGGATTTGGCAAA	56077
Qy	4724	CTGATTTTACCCACAAGCCCCAACTCTCAGGGATTTTCAGTATCTACTAGTCTGGGTARATA	4783
Db	56076	TTGACTTTACTCAGATGCCCCGAGTCAGGAACCTAAATACCTCTTGTCTGGGTAGACA	56017
Qy	4784	CTTTTCACGGTTGGGCARAGGCCCTCCCTGTATAGACAGAAAGGCCCAAGAGTAATAAA	4843
Db	56016	CTTTTCACTGGATGGGTAGAGGCCCTTTCCACAGAGGCTCTGAGAAGGCCACCGTGTCAITTT	55957

QY	4844	AGGCATAGTTCATGAATAATTTCCAGATTCGGACCTTCCCGAGGCTTACAGATGACA	4903
DB	5956	CTTCCCTTCTGTGAGACATAAATCTCTCAGTTCCTGGCTTCCACCTCTATACAGTCCGATA	55897
QY	4904	ATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCACAGCGTTCAGGTATACGAT	4963
DB	55896	ACGGATGAGCTTTACTAGTTAAATTTACCAAGCAGTTCCTCAGGCTCTTGGTATTCAGT	55837
QY	4964	ATCATTACACTCTGGCTGAAAGGCCACAGTCCTCAGGGGAAGTFCGAGAAAATGAATGAAA	5023
DB	55836	GGAACTTCAATTCCTTACCATCTCTTCAAGAAAGGTAGAAATAGACTAATGGTC	55777
QY	5024	YACTCAAGAGACATCTAAAAAGCAACCCAGGAAACCCACCTCACATGGCTGTCGTCT	5083
DB	55776	TTTTAAAGACACACTCTACCAAGCTCAGCTCTCAATTTTAAAGGACTGGACAGTACTTT	55717
QY	5084	TGCTATAGCTTTAAAAAGAATCTGCAACTTTTCCCAAAAAAG---CAGGACTTTAGCCCAT	5140
DB	55716	TACCTCTTGCCCTTCTCAGAAATTAGAGCCTGTCTCGAGATGCTACAGGGTACAGTCCAT	55657
QY	5141	ACGAATGCTGTATGAAAGGCCCTTCATTA---ACCAATGACCTTGTGCTTGACCCCAAGAC	5197
DB	55656	TTGAACCTTTTATACGACGACACTTTCTTACTCAGGCCCAACCTTGTCTCCGACACACAGCC	55597
QY	5198	AGCCAACCTAGTTGCAGACATCACCTCTCTTACGCCAAATATCAACAAGTTCTTAAAAACATT	5257
DB	55596	CTCTAGGTGACTATCTCTCAGTCTCTCCAGCAGGCTA--GACAGGAAATTCGCCAGGCTGC	55539
QY	5258	ACAAGGAACTATCTCCCTGAGAAGAGGGGAAAGAACTATTTCACCCGWTGATCATGTATT	5317
DB	55538	TAATCTTCTCTTGCCCTACTCCAGATTCTCAACCATATGAACACACCCCTAGCTGGACAATC	55479
QY	5318	AGTCAAGTCCCTTCYCTCTAAATTTCCCATCCCTAGATACATCCTCGGGAAGGACCTTACCC	5377
DB	55478	AGTCTCTGTTAAGAAATCTGACCCCTCAAACTCTACAACCTCGATGGACTGGACCCCTACTT	55419
QY	5378	AGTCATTTTATTTACCCCAACTGCCGTTTAAAGTGGCTCGAGTGGAGTCTTGGATACATCA	5437
DB	55418	AGTCATCTATAGTACCCCAACTGCCGTTCACCTTCGAGGATCTCTCCCACTGGGTTTACCG	55359
QY	5438	CACCTGAGTCAA	5449
DB	55358	TTCCAGAGTAAA	55347
RESULT 3			
US-10-528-659-2/c			
; Sequence 2, Application US/10528659			
; Publication No. US2006009594A1			
; GENERAL INFORMATION:			
; APPLICANT: NAGOYA INDUSTRIAL SCIENCE RESEARCH INSTITUTE			
; APPLICANT: GIFU INTERNATIONAL INSTITUTE OF BIOTECHNOLOGY			
; APPLICANT: YAMADA, Yoshihiko			
; APPLICANT: YOKOTA, Mituhiko			
; TITLE OF INVENTION: Method for diagnosing a risk of hypertension			
; FILE REFERENCE: C0200701			
; CURRENT APPLICATION NUMBER: US/10/528,659			
; PRIORITY FILING DATE: 2005-03-22			
; PRIOR APPLICATION NUMBER: JP P2002-280034			
; PRIORITY FILING DATE: 2002-09-25			
; NUMBER OF SEQ ID NOS: 36			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 2			
; LENGTH: 50000			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-528-659-2			

Query Match 1.8%; Score 132.2; DB 6; Length 50000;  
Best Local Similarity 44.2%; Pred. No. 1.2e-28;  
Matches 730; Conservative 71; Mismatches 797; Indels 54; Gaps 11;  
Qy 2706 GTTGTCCTTTTATACCAGCTGTACTAGCCCTTATCTGTGMYTTCCTCAATAACGAG 2765

Db	18660	GTCTTGCTATCCACTCTGTGGTGC	CAAAACCATATACTCTCTATCTCTAATACTCC	18601
Qy	2766	GAAGCAGATGGTTTACASTCTGGACCT	TTMAGGATGCTTCTTGCACTCCCTGTACAT	2825
Db	18600	TCCACAACCATATTTCTGTCTGGATCT	TGAAGATGCTTCTTACCTTTTGCAC	18541
Qy	2826	CTGACTCTCAATTCCTGTTTGCTT	-----TGAAGATACTTCAAAACCCARCATCT	2876
Db	18540	TCCTCACCCAGCTCTCTTCACTTACAT	TGGACTGACCCCTGACACCCATCAGTCTCNG	18481
Qy	2877	CAACTCACCTGGACTRTTTTACCCCAAGG	TTTACGGGATGAGYCCCATCTATTTTGGCCAG	2936
Db	18480	CAACTTACCTGGCTGTACTGCCACAAGCT	TTCAGGCAGCTCCATTTACTTCAGTCAA	18421
Qy	2937	GCATTAGCCAGACATTGAGYCA	RTTWTCACTTGGACACTCTTGTCTTCRGTAKGTG	2996
Db	18420	GCCTTTTCTCATAAATTTACTTTTCTGT	CTCAT---CTGCTGTGACCTTTATTTCAATATTTT	18364
Qy	2997	GATGATTTACTTTTTCGCGCTTTCAGAAA	ACCTTGTGCATCAAGCACCAAGCCTC	3056
Db	18363	GATGGCTTCTACTTTATAGCCCTCTAT	GAATCTTCCACAGCACCCCTCTGCTC	18304
Qy	3057	TTMAATTTCTCGYACCTGTGGCTACAWG	TTTTTCCAAACSSARAGCTCACTCTGTCTCA	3116
Db	18303	CTCCAAACATCTATTTCAAGGGAATAT	CTGTATCCCCCTCAAAGCCAAATTTCTTC	18244
Qy	3117	CAGCAGTTAAATACTTAGGCTTA	TAARATTTCAAAAGCACCGCCCTCAGTGAGAA	3176
Db	18243	TCATCCGTTTACCTTA--TCGGCATAA	TTCTTCAATAAAATACAAGTGCTCTCCCTCTGAT	18186
Qy	3177	YRYATCCAGCTATACTGGCTTATCTGCT	ATCYC-AAAACCTTAAAGCAACTAAGRRTT	3235
Db	18185	CTGTGTCGGCTTAATCTCCAAACCCCA	ATCCCTTCTCAAAAGCAAAATCTCTTCTCTT	18126
Qy	3236	CTTGGCTTAAYAGGYTTCTGCGAAWAT	GGATTTCCCAAGGTWTGGCRAAATAGCCAGY	3295
Db	18125	CCTAGACATGTTGGGTACTTTCACCT	TTTGGTA-TACCTGGTTTGGCATCTGACTAAAC	18067
Qy	3296	CATTAAATACSTAAATTAAGGAAACT	TCAGAAAGCCAAATACCCATTTARTAGATGGAVAM	3355
Db	18066	CATTACATAACTCACAAAAGGAAAC	CTAGCTGACCCCA-----TAGATCTTAAA	18017
Qy	3356	CTGAAGYMRAGTGGCTTTCCAGGCCCT	TAAAGAGGCTTAAACCCAGYCCGAGTGT	3415
Db	18016	TCCTTTCCCACTCCCTCTTATGTCT	CTTTAAACAGCCCTAGAGCTACTCCCACTA	17957
Qy	3416	AAGYTTGCCAACRGCGCAAGACTTTT	STTATYATRTCAAGAAAAAACAGRAAYAGCTC	3475
Db	17956	GCCTCCCTAACTCATCCCAACCTTTT	-----CATTACACAGCCAAAGTACAAGCT	17902
Qy	3476	TRGAGTCTTTACACAGTCCRAGGAY	GAGCTTGCAACCYRGTGGRYACCTGASTAAGG	3535
Db	17901	GTGCGGTCAGAATCCTTACAAAGAGCT	TGGACAGTGCCCTCGACGCTTTCTGTCTAAAC	17842
Qy	3536	AAAYTGATGTAGTGGCAAGGGTTC	RYCATTTATYGGGTAGTGGTGCGAGTAGCAG	3595
Db	17841	AGATTGACCTTACTGTTTAGGCTG	GGCTCTCATGTCTGTGTGGGAGCTGCCACTGCTT	17782
Qy	3596	TYKTAGTATCTGAAGCAGTTTAAATA	ATAATACAGGGRAGAGATCTTACTGTGTGGACATCTC	3655
Db	17781	TAATACTTTTAAAGGCCCTCAANT	CACAACTATATGATCACTCTCTACAGTTCTC	17722
Qy	3656	ATGAKGTGAAYRGATACATCACTGC	-----TAAAGAGACTTTGTGGCTGT	3700
Db	17721	ATAACTTCAAAATGATATTTTCTCT	CTCAACCTGATGATATACTTTCTGCGCTCTCC	17662
Qy	3701	CAGACACYTTTACTTAAATRTCAGGCT	CTATTACTGAAGGCCAGTGTGCGRACTGT	3760
Db	17661	CCCTCCCGCGCTCCTTCAGTTGA	CTCACTCTTCTTGTGAGTCTCCCAAAATACCATTG	17602
Qy	3761	GCATTGTGCAACTCTTAACCCAGYCN	CAATTTCTCCAGACAAATGAAGAAAAATARAAY	3820

Db	17601	TTCTCTGGCTGGACTTCAATCTGGCTCCCAATTAATTCCTGATACCAACCTGATCCCC	17542
Qy	3821	ATAACTGTCAACAARATAATTTCTCAAACTATGCCACTCGAGGGACCTTTTAGAGTTTC	3880
Db	17541	ATGACTGTATCTCTCTGATCCACCTAGCAATTCACATTTTCCCATTAATTCCTTCTTTTC	17482
Qy	3881	CYTTGACTGATCCYGACCTTCAACTTGTATCTATGATGGAAGTTCCTTTGTAGAAA----	3935
Db	17481	CTGTTCTCTCACTCTGATC-ACACTTGGTTTATTGATGCGAGTTCCTCCTAGCCTAATCAC	17423
Qy	3936	-AAGGACTTCGAAAAGYGGGTATGCGAGTGGTCACTGATATATGGAAATAYTTGAAAGTAAT	3994
Db	17422	CACTCACGACGAAGCAGGCTACGCTATAGTATCTTTTTCACATCTATCATTTGAGGCTACC	17363
Qy	3995	CCCTCAGCTCCAGGAACCTAGTCTYAGCTGCGAGAACTAAATAGCCYTCAYTKGGSCACTA	4054
Db	17362	ACTCTGCCCACTCCACTCCCTCTCAGCAAGCCGAACCTCATTTGCCTTAATCGAGCCCTC	17303
Qy	4055	GAATTAGGAGAAAGRAAAAAGGYYAAATATATATACAGACTCTTRATATGCTYACCTAGTC	4114
Db	17302	ACTCTTGCAAAAGGATTGTGTCTCAGTATTATATCTGACTCTAAATATGCTTCCATATT	17243
Qy	4115	NTCCATGCCCATGMRGCAATATGSPAGAAAGGAATTCCTTAACCTTCYVAGRGAAACACT	4174
Db	17242	CTCCAGCTCCATTTCTGTTATATGGCTGAAAGAGGTTTCTCACTATACGAGGGTCTCTCC	17183
Qy	4175	ATCAMACATCNGMAAGCCATTAGGARATTTATTAYTGGCGWGTACAGAAACCTTARAGAGGTG	4234
Db	17182	ATTATTAAATGGCTTCTTCATAAAAACTCTTCTTAAAGCTGCTCTACTTCCAAGGAAGCTA	17123
Qy	4235	GMAGTCTTACACTGCVGGGGTCAATCANAAAAGGAAAGGGAATATASAAAGRGAAATGCG	4294
Db	17122	G-CGTCAITTCATGCAAGGCCATCAAAGGCATCAGATCCATCCTCAGGCGCAAGCT	17064
Qy	4295	CAAGCAKATATTGAAGCMMAAAGAGCTGCAAG	4326
Db	17063	TATGCTGTATAAGGTAGCTTAAGAGGACGAGCTAG	17032

## RESULT 4

US-10-539-228-703/c  
; Sequence 703, Application US/10539228  
; Publication No. US20060154250A1

```

, GENERAL INFORMATION:
, APPLICANT: David W. Morris
, APPLICANT: Marc S. Malandro
, TITLE OF INVENTION: Novel Compositions and Methods in Cancer
, FILE REFERENCE: CHIR0052-101 (PP023370.0003)
, CURRENT APPLICATION NUMBER: US/10/539,228
, CURRENT FILING DATE: 2005-06-17
, PRIOR APPLICATION NUMBER: US 10/322,281
, PRIOR FILING DATE: 2002-12-18
, NUMBER OF SEQ ID NOS: 866
, SOFTWARE: FastSeq for Windows Version 4.0
, SEQ ID NO 703

```

; LENGTH: 88493

LENGTH: 00455  
TYPE: DNA

ORGANISM: Mus musculus

CREATING: **FEATURE:**

```

FEATURE:
NAME/KEY: misc feature

```

; NAME/KEY: WISC\_reacure  
: LOCATION: (1) (88493); LOCATION: (1) : : (88433)  
; OTHER INFORMATION: n = A-T-C or G; OTHER INFORMATION  
IIS-10-539-228-703

Query Match	0 q%	Score	66.2	DB	6	Length	88493
-------------	------	-------	------	----	---	--------	-------

Query Match	0.9%	Score 66.2;	DB 6;
Best Local Similarity	51.3%	Pred No 5	5e-08;

Best Local Similarity 51.3%; Pred. No. 5.5e-08;  
Matches 220: Conservative 4: Mismatches 197

Matches 220; Conservative 4; Mismatches 197; Indels 8; Gaps 3;

[illegible]

4738 AAGCCAAACCTCAGGGATTTCAGTATCTACTAGTCTGGGTAAATCTTCACGGGTGG 4797

[illegible]

b 66709 AAACCAGAAAAATTTAGATATAAATATTTACTGATGTTTACAGATACCTTTTCAGGTGG 6665

**A708 CCCCCCCCCCCCCCCCTTCCCTTTTCATTA**

4798 GCARAGGCCTTCCCCCTGTAGGACAGAAAAGGCCCAAGAGGTAATAAAGGCACCTAGTTCAT 4857



Db 66649 ACTGAGGCTTTCCCAACCAATGACACCATTCATAGTGACAAAGAGCTTTAGAG 66590  
Qy 4858 GAATATATCCAGATTCCGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTC 4917  
Db 66589 GACATAGTGCACAGGTATGGGCTTCTCTCA--GATGATAGTCAGACAAATGGACCACTCTT 66532  
Qy 4918 CAGGCCACAGTAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGC 4977  
Db 66531 GTGCCCCAGTGAAGTCAGAGACTAGTCAATATCTCGAGATGATTAATAACT----TGT 66476  
Qy 4978 GCTTGAAGGCCAGTCTCTCAGGAAAGTTCGAGAAATGAATGAAYACTCAAGGACAT 5037  
Db 66475 ATATATAGACCCAGAGTTCAGGACAGGTAGAAAAATAAACAAATCTTAAGAGACC-- 66418  
Qy 5038 CTAATAAAGCAACCCAGGAACCCACCTCACATGGCTGCTTGTTCCTATAGCCCTTA 5097  
Db 66417 TTAACCTAACTGACCTAAGAGACTAGTGAATGACTGAATAACTCTCTTCCCTTTACTCTC 66358  
Qy 5098 AAAAGAATCTGCAACTTTTCCCAAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGA 5157  
Db 66357 TATAGGCGTGAATCTCCCTTACCAGATGGAATCTACCCCTTTTAAATAATATGTTGCGG 66298  
Qy 5158 AGGCCCTTC 5166  
Db 66297 ATACCACTC 66289

RESULT 5

US-11-218-305-4196  
; Sequence 4196, Application US/11218305  
; Publication No. US20060141495A1  
; GENERAL INFORMATION:

; APPLICANT: MONSANTO TECHNOLOGY, LLC  
; APPLICANT: Mc Laird, Paul L.  
; APPLICANT: Tao, Nengbing  
; APPLICANT: Wu, Kunheng  
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping  
; TITLE OF INVENTION: Corn.

; FILE REFERENCE: 38-21 (53660) B  
; CURRENT APPLICATION NUMBER: US/11/218,305  
; CURRENT FILING DATE: 2005-09-01  
; PRIOR APPLICATION NUMBER: US 60/606,880  
; PRIOR FILING DATE: 2004-09-01  
; NUMBER OF SEQ ID NOS: 25043  
; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 4196  
; LENGTH: 2579  
; TYPE: DNA  
; ORGANISM: Zea mays

US-11-218-305-4196

Query Match 0.6%; Score 48.4; DB 7; Length 2579;  
Best Local Similarity 45.5%; Pred. No. 0.0012;  
Matches 160; Conservative 3; Mismatches 189; Indels 0; Gaps 0;

Qy 4746 ACTCAGGGATTTCAGTATCTACTAGTCTGGGTATATCTTTCACGGGTTGGCAGAGGC 4805  
Db 1378 ATCAAGGGGCATCGGTGTGTGTAGTTCCTGACTACTTCCACCAATGGACTGAGC 1437  
Qy 4806 CTTCCCTCTAGACAGAAAGCCCAAGAGGTAAATAGGCACTAGTTCATGAATAAT 4865  
Db 1438 CGTGTCTTAAGAACATGACACACAGGAGGTAATTGAGTTTATACTGAGCATATTAT 1497  
Qy 4866 TCCAGATTCCGACTTCCCGAGGCTTACAGATGACATAGCCCTTCCAGGCCAC 4925  
Db 1498 TCATAGATTCCGCAATTCGCCAGACCTTAATCAAGATCAGGACTTCTTTATGTCAA 1557  
Qy 4926 AGTAACCCAGGGAGTATCCAGCGGTTAGGTATACGATATCACTTACATCGCCCTGAAG 4985  
Db 1558 GGAGGTACGTGAATTTGCTGAATTTATACAGAAATTAAGCTCTTAATTCATCTCATTA 1617  
Qy 4986 GCCACAGTCTCAGGAGAGTCAGAAATGAATGAAYACTCAAGAGGACATCTAAAAA 5045

Db 1618 TGCTCAGGCCAATGGACAGGCCGAGTCTAGTAATAGGACATGATTAATTTGATAAAAA 1677  
Qy 5046 GCAAAACCCAGAAACCCACCTCACATGGCCTGTCTGTTCCTATAGCCTTA 5097  
Db 1678 GAAGATATCTGATAATCTTAACATTGGCATAAGATTTTCTCTGAAGCTTTA 1729

RESULT 6

US-10-539-228-645/c  
; Sequence 645, Application US/10539228  
; Publication No. US20060154250A1  
; GENERAL INFORMATION:

; APPLICANT: David W. Morris  
; APPLICANT: Marc S. Malandro  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
; FILE REFERENCE: CHIR0052-101 (PP023370.0003)  
; CURRENT APPLICATION NUMBER: US/10/539,228  
; CURRENT FILING DATE: 2005-06-17  
; PRIOR APPLICATION NUMBER: US 10/322,281  
; PRIOR FILING DATE: 2002-12-18  
; NUMBER OF SEQ ID NOS: 866

; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 645  
; LENGTH: 61791  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:

; NAME/KEY: misc feature  
; LOCATION: (1)..(61791)  
; OTHER INFORMATION: n = A,T,C or G

US-10-539-228-645

Query Match 0.6%; Score 48.4; DB 6; Length 61791;  
Best Local Similarity 51.2%; Pred. No. 0.014;  
Matches 125; Conservative 5; Mismatches 111; Indels 3; Gaps 1;

Qy 6614 AGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGGTCCCGACTCCC 6673  
Db 58689 AGAATTACAGAACCATATATACATTGCACTAGATAGCAATATAGAGAAATAGACAACTCA 58630  
Qy 6674 TGCTCACCCTGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGCTCTTCAAAATCGAGAG 6733  
Db 58629 TTAATGCTCTACAAGATTCCCTTTCTCTGCTGAAACAGTTAAACCACTTGAAGAG 58570  
Qy 6734 CTTTAGACTYGTCTAACCGCTGARAGAGGGGACCTGTTTATTTTATTTAGGGGAAGATGCT 6793  
Db 58569 GATTGGAGTTGTT--TCTTCAGCAGGGTGGTTTATGTTTTCAGCTCTTGGAGAAAGATGCT 58513  
Qy 6794 GTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAATTCAGATCGAA 6853  
Db 58512 GTGCTTATGCTGCTCACTCCAAAGTAGTCAAAAGAGATATGGCTCAGCTTAGAAAAAGAT 58453  
Qy 6854 TACA 6857  
Db 58452 TAAA 58449

RESULT 7

US-10-554-711-202  
; Sequence 202, Application US/10554711  
; Publication No. US20060115806A1  
; GENERAL INFORMATION:

; APPLICANT: University of Georgia Research Foundation  
; APPLICANT: McDonald, John F.  
; TITLE OF INVENTION: Global Analysis of Transposable Elements  
; FILE REFERENCE: as Molecular Markers of Cancer  
; FILE REFERENCE: 21099.0075P1  
; CURRENT APPLICATION NUMBER: US/10/554,711  
; CURRENT FILING DATE: 2005-10-28  
; PRIOR APPLICATION NUMBER: 60/466,798  
; PRIOR FILING DATE: 2003-04-29  
; NUMBER OF SEQ ID NOS: 778

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 202
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/note =
; OTHER INFORMATION: Synthetic Construct
US-10-554-711-202

Query Match
Best Local Similarity 0.6%; Score 42.6; DB 6; Length 50;
Matches 45; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7127 AAGGCAACCCCTCTCAGGAAATCTCAGCTGCACAACTTACTAGGCC 7175
      |||||
Db 2 AAGGCAACCCCTCCGAGGAAATCTCACTGCACGACCCCTACTAGGCC 50

RESULT 8
US-10-517-441-440/c
; Sequence 440, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
; APPLICANT: FOEKENS, John
; APPLICANT: HARBECK, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime P.
; APPLICANT: MARX, Almuth
; APPLICANT: HOEFER, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; PRIOR FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 440
; LENGTH: 6175
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chemically treated genomic DNA (Homo sapiens)
US-10-517-441-440

Query Match
Best Local Similarity 0.6%; Score 42.2; DB 6; Length 6175;
Matches 98; Conservative 14; Mismatches 107; Indels 0; Gaps 0;

QY 3198 TATCTCTCATCYCAAAACCCCTAAAGCACTAAGRGRTTCCTTGGCRTAAYAGGYTTCTGC 3257
      |||||
Db 5675 TTCTTAACCACTACCAAAACGCAAAACCTTCTATAATAATAATTACTTCTTC 5616

QY 3258 CGAAATGATTTCCCGAGTWTGGCRAATAGCCAGGYCATTAATACASTAATTAAGGA 3317
      |||||
Db 5615 TTATAAAAAACCCATTAATAATACCTTATAATCAAAACCTTAATAAAACAATACTTAAATA 5556

QY 3318 AACTCAGAAGCAATACCATTARTTAAGATCGAYAMCTGAAGYVRAAGTGGCTTCCA 3377
      |||||
```

```
Db 5555 TAATCAACACACAAAAAACTATTCAAACACTATATAATACTAAATCCAAACTCGTATCACT 5496
QY 3378 GGCCCTTAAGAGGCTTTAAACCCCAAGYCCAGTGTTA 3416
      |||||
Db 5495 TACCACACACACCAACCAATAAATCTAAAAACCAAAATATTA 5457

RESULT 9
US-11-218-305-4195
; Sequence 4195, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLaird, Paul L.
; APPLICANT: Tao, Kunsheng
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; TITLE OF INVENTION: Corn.
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; PRIOR FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4195
; LENGTH: 3016
; TYPE: DNA
; ORGANISM: Zea mays
US-11-218-305-4195

Query Match
Best Local Similarity 0.6%; Score 42; DB 7; Length 3016;
Matches 156; Conservative 3; Mismatches 193; Indels 0; Gaps 0;

QY 4746 ACCTCAGGATTTTCAGTATCTACTGTCTGGTARATACTTTTCAGGGTTGGGCARAGC 4805
      |||||
Db 1998 ATCAAGGGCATCGTTGCTGTAGTTGCCACTGACTACTTACCAATGAGCTGAAGC 2057
      |||||
QY 4806 CTTCCCTGTAGGACAGAAAGCCCAAGAGGTATAAAGGCATAGTTTCATGAATAAT 4865
      |||||
Db 2058 CGTTGCTCTGAAGAACATGACACACAAGGAGTAAATGAGTTCAATACTGAGCATATTAT 2117
      |||||
QY 4866 TCCAGATTCGACTTCCCGAGGCTTACAGAGTACAGATACCTGCTTCCAGGCCAC 4925
      |||||
Db 2118 TCATAGATTTCGGCATTTCCCGAGCCTTCACACTACAGATCAAGGAGCTTTTATGTCAA 2177
      |||||
QY 4926 AGTAACCCAGGAGTATCCAGGCTTAGGTATACGATATACACTGCGCCTGAAG 4985
      |||||
Db 2178 GGAGTACGTGAATTTGCTGAATTAACAGAAATTAAGCTGCTTAATTCATCTCCATATTA 2237
      |||||
QY 4986 GCCACAGTCTCAGGAAAGGTCGAGAAATGAATGAAAYACTCAAGAGGACATCTAAAAA 5045
      |||||
Db 2238 TGCTCAGGCAATGACAGGCGAGCTAGTAGTATAGGACATTGATTAAATTTGATAAAAA 2297
      |||||
QY 5046 GCAACCCAGGAAACCCACCTCACATGGCTGCTGTTGCTGCTATAGCCTTA 5097
      |||||
Db 2298 GAAGATATCTGATAATCTCAAACATTGGCATAAGATTTTGTCCGAGGCTTTA 2349
      |||||

RESULT 10
US-10-554-711-192
; Sequence 192, Application US/10554711
; Publication No. US20060115806A1
; GENERAL INFORMATION:
; APPLICANT: University of Georgia Research Foundation
; APPLICANT: McDonald, John F.
; TITLE OF INVENTION: Global Analysis of Transposable Elements
; TITLE OF INVENTION: as Molecular Markers of Cancer
; FILE REFERENCE: 21099.0075P1
; CURRENT APPLICATION NUMBER: US/10/554,711
; CURRENT FILING DATE: 2005-10-28
; PRIOR APPLICATION NUMBER: 60/466,798
```

Qy 3318 AACTCAGAAAGCCCAATAACCCATTATTAAAGATGGAYAMCTGAAGYMYRAAGTGGCTTTCCA 3377  
 Db 5555 TAATCAACAACAAAAAACTATTCAAACCTATAATAAATACTAAATCAAACCTCATATCACT 5496  
 Qy 3378 GGCCCTTAAGAAGGCCCTTAACCCCAAGYCCCAAGTGTTA 3416  
 Db 5495 TACCAACAACAACCAATAAATCTAAAAACAAAAATATTA 5457

RESULT 12  
 US-11-256-173-12  
 ; Sequence 12, Application US/11256173  
 ; Publication No. US20060140979A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: University of Sheffield  
 ; TITLE OF INVENTION: Antigenic Peptides  
 ; FILE REFERENCE: toxin  
 ; CURRENT APPLICATION NUMBER: US/11/256,173  
 ; CURRENT FILING DATE: 2005-10-24  
 ; PRIOR APPLICATION NUMBER: US/10/311,879  
 ; PRIOR FILING DATE: 2003-03-18  
 ; NUMBER OF SEQ ID NOS: 32  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 12  
 ; LENGTH: 7975  
 ; TYPE: DNA  
 ; ORGANISM: Staphylococcus aureus  
 US-11-256-173-12

Query Match 0.5%; Score 40.2; DB 7; Length 7975;  
 Best Local Similarity 46.0%; Pred. No. 0.98;  
 Matches 93; Conservative 10; Mismatches 99; Indels 0; Gaps 0;

Qy 3970 TGATAATGGAATAYTTGAAGTAGTAATCCCTCACTCCAGGAACCTAGTCTYAGCTRCAGA 4029  
 Db 7089 TAATAATGCGAACAATGCATTAAATGGTAGCGAAAACCTTAAACAATGCGAAAACAAGCAGC 7148  
 Qy 4030 ACTAATAGCCYTCAYTKGGGCACCTAGAAATTAGGAGAAGRAAAAAGGGYAAATATATATAC 4089  
 Db 7149 GATTACAGCAATCAATGGCGCATCTGATTTAATCAAAAACAAGAAGATGCATTAAAGC 7208  
 Qy 4090 AGACTCTRTARTAGCTYACCTATGTCNTCCATGCCCAATGMRGCAATATGSAAGAAAGGGA 4149  
 Db 7209 ACAAGCTAATGGTGTCTCAACGGCTATCTAATGACAAAGATGTACAGCACAATGCGACTGA 7268  
 Qy 4150 ATTCCTAACTTCYGAGRGACA 4171  
 Db 7269 ACTGAACACGGCAATGGGCACA 7290

RESULT 13  
 US-10-471-571A-4495  
 ; Sequence 4495, Application US/10471571A  
 ; Publication No. US20060115490A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CHIRON SPA  
 ; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS  
 ; FILE REFERENCE: P026927WO  
 ; CURRENT APPLICATION NUMBER: US/10/471,571A  
 ; CURRENT FILING DATE: 2003-09-12  
 ; PRIOR APPLICATION NUMBER: GB-0107661.1  
 ; PRIOR FILING DATE: 2001-03-27  
 ; NUMBER OF SEQ ID NOS: 5642  
 ; SOFTWARE: Seqwin99, version 1.03  
 ; SEQ ID NO 4495  
 ; LENGTH: 28605  
 ; TYPE: DNA  
 ; ORGANISM: Staphylococcus aureus  
 US-10-471-571A-4495

Query Match 0.5%; Score 40.2; DB 6; Length 28605;  
 Best Local Similarity 46.0%; Pred. No. 2.6;  
 Matches 93; Conservative 10; Mismatches 99; Indels 0; Gaps 0;



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2006, 17:10:51 ; Search time 30582 Seconds  
(without alignments)  
13863.723 Million cell updates/sec

Title: US-10-717-580-11

Perfect score: 7489.6

Sequence: 1 caacaatcggtataataacc.....tattaataatctgcactgtg 7582

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: \*  
1: gb\_est1: \*  
2: gb\_est3: \*  
3: gb\_est4: \*  
4: gb\_est5: \*  
5: gb\_est6: \*  
6: gb\_hic: \*  
7: gb\_est2: \*  
8: gb\_est7: \*  
9: gb\_est8: \*  
10: gb\_est9: \*  
11: gb\_gss1: \*  
12: gb\_gss2: \*  
13: gb\_gss3: \*  
14: gb\_gss4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2180.4	29.1	2748	6	CR605851
2	2176.4	29.1	2749	6	CR617248
3	2174.4	29.0	2500	6	CR622175
4	2169.4	29.0	2748	6	CR625046
5	2147.4	28.7	2716	6	CR613169
6	969	12.9	1071	4	BX365086
7	921.2	12.3	998	4	BX337769
8	909.6	12.1	1058	4	BX378303
9	908.6	12.1	935	4	BX391741
10	885.2	11.8	955	4	BX347952
11	885	11.8	1019	4	BX439636
12	866.6	11.6	1014	4	BX378304
13	848.8	11.3	6400	6	CR749521
14	834.8	11.1	924	4	BX347314
15	820.8	11.0	931	4	BX389656
16	820	10.9	903	4	BX408752
17	812	10.8	872	4	BX326647
18	810.4	10.8	6427	6	CR857090
19	804	10.7	883	4	BX408733

c 20	803.8	10.7	966	4	BX380176	BX380176 BX380176
c 21	789.2	10.5	877	4	BX347619	BX347619 BX347619
c 22	776.2	10.4	921	4	BX408734	BX408734 BX408734
c 23	771	10.3	836	4	BX388766	BX388766 BX388766
c 24	768.6	10.3	995	4	BX389657	BX389657 BX389657
c 25	767.6	10.2	808	9	DA834147	DA834147 DA834147
c 26	765.4	10.2	844	4	BX347111	BX347111 BX347111
c 27	764.8	10.2	828	4	BX367907	BX367907 BX367907
c 28	753.8	10.1	846	4	BX368078	BX368078 BX368078
c 29	751.8	10.0	820	9	DA833993	DA833993 DA833993
c 30	751.6	10.0	900	4	BX459153	BX459153 BX459153
c 31	734.2	9.8	836	13	C2457645	C2457645 MCF745b17
c 32	732.6	9.8	770	9	DR004257	DR004257 TC114345
c 33	724.6	9.7	758	9	DA833923	DA833923 DA833923
c 34	723.4	9.7	769	4	BX357208	BX357208 BX357208
c 35	723.4	9.7	879	4	BX325902	BX325902 BX325902
c 36	719.4	9.6	924	4	BX409328	BX409328 BX409328
c 37	718.6	9.6	761	9	DA832031	DA832031 DA832031
c 38	715	9.5	782	9	DA833912	DA833912 DA833912
c 39	709.4	9.5	771	2	BI087886	BI087886 602852690
c 40	704.2	9.4	792	9	DA830020	DA830020 DA830020
c 41	700.4	9.4	840	9	DA834086	DA834086 DA834086
c 42	697.2	9.3	721	1	AU138405	AU138405 AU138405
c 43	686.2	9.2	819	9	DA830049	DA830049 DA830049
c 44	679.4	9.1	799	13	C2450650	C2450650 MCF731h04
c 45	678.2	9.1	832	14	DU836611	DU836611 MUGQ_CH25

## ALIGNMENTS

CR605851 2748 bp mRNA linear HTC 21-JUL-2004  
full-length cDNA clone CS0DE012YJ24 of Placenta of Homo sapiens (human).  
ACCESSION CR605851  
VERSION CR605851.1 GI:50486658  
KEYWORDS HTC; CNSLT cDNA.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 2748)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue  
REFERENCE 2 (bases 1 to 2748)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.  
FEATURES  
Location/Qualifiers  
source  
1..2748  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DE012YJ24"  
/tissue\_type="Placenta"  
/plasmid="pCMVSPORT\_6"  
ORIGIN  
Query Match 29.1%; Score 2180.4; DB 6; Length 2748;

Best Local Similarity 98.4%; Pred. No. 0;			
Matches 2171; Conservative 34; Mismatches 0; Indels 1; Gaps 1;			
QY	5352	GATACATCTCGGAAGAGCCCTACCCAGTCAATTTATYATACCCCAACTGCGGTAAAGTG	5411
Db	544	GATACATCTCGGAAGAGCCCTACCCAGTCAATTTATCTACCCCAACTGCGGTAAAGTG	603
QY	5412	GCTGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAATCTGGATCTGCAAGGAA	5471
Db	604	GCTGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAATCTGGATCTGCAAGGAA	663
QY	5472	CCTGAAATCCAGGAGACAAACCTAGCTATCTCTGTGAACCTCTAGAGGATTTGCGCTG	5531
Db	664	CCTGAAATCCAGGAGACAAACCTAGCTATCTCTGTGAACCTCTAGAGGATTTGCGCTG	723
QY	5532	CTCTTCAAACAACACGAGGAGAAAGTAACATAAATCAATAATCCCATGGSCTCCC	5591
Db	724	CTCTTCAAACAACACGAGGAGAAAGTAACATAAATCAATAATCCCATGGSCTCCC	782
QY	5592	TTATCATATTTCTCTKSTASTGTSTTTTACCTSTTTTCACTCTCACTGCAACCCCTCC	5651
Db	783	TTATCATATTTCTCTKSTASTGTSTTTTACCTSTTTTCACTCTCACTGCAACCCCTCC	842
QY	5652	ATGCGCGCTGATGACAGTAGTCCCTVACCMAGAGTTTCTATGGAGAAATGCAGGTCC	5711
Db	843	ATGCGCGCTGATGACAGTAGTCCCTTACCAGAGTTTCTATGGAGAAATGCAGGTCC	902
QY	5712	CGGAATATTTGATGCCCATCGTATAGGAGTCTTTSTAAGGAAACCCCACTCACTGC	5771
Db	903	CGGAATATTTGATGCCCATCGTATAGGAGTCTTTSTAAGGAAACCCCACTCACTGC	962
QY	5772	CCACACCATATGCCCGCAATGCTATCACTCTGCCACTTTTGCATGCGATGCAATAC	5831
Db	963	CCACACCATATGCCCGCAATGCTATCACTCTGCCACTTTTGCATGCGATGCAATAC	1022
QY	5832	TCATTAATGGACGAGAAATGATTAATCTAGTTGTCTGAGGACTTTGAGTCACTGT	5891
Db	1023	TCATTAATGGACGAGAAATGATTAATCTAGTTGTCTGAGGACTTTGAGTCACTGT	1082
QY	5892	CTGTTGGACTTACTTCAACCAACTGGTATGTCTGATGGGGTGGAGTTCAAGATCAGGC	5951
Db	1083	CTGTTGGACTTACTTCAACCAACTGGTATGTCTGATGGGGTGGAGTTCAAGATCAGGC	1142
QY	5952	AGAGAAAAACATGTAAAGAGATATCTCCCACTCACCSGGGTACATGGCACTCTAG	6011
Db	1143	AGAGAAAAACATGTAAAGAGATATCTCCCACTCACCSGGGTACATGGCACTCTAG	1202
QY	6012	CCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCTCCGTACCCATCTCGCT	6071
Db	1203	CCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCTCCGTACCCATCTCGCT	1262
QY	6072	GGTAAGCCTATTAAATACACCTCTACGGGTCCATGAGTCTCGGCCCAAAACCCCTAC	6131
Db	1263	GGTAAGCCTATTAAATACACCTCTACGGGTCCATGAGTCTCGGCCCAAAACCCCTAC	1322
QY	6132	TAACTGTTGGATATGCTCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACTGA	6191
Db	1323	TAACTGTTGGATATGCTCCCTGAACTTCARGCCATATGTTTCAATCCCTGTACTGA	1382
QY	6192	ACAAATGGAACAATTCAGACAGAAATAAACCACTTCCTGTTTGTAGTAGGACCTTGT	6251
Db	1383	ACAAATGGAACAATTCAGACAGAAATAAACCACTTCCTGTTTGTAGTAGGACCTTGT	1442
QY	6252	TTCCAAATSTGGAATAACCATACCTCAAACTCTGTTGTAATTTAGCAATCTAC	6311
Db	1443	TTCCAAATSTGGAATAACCATACCTCAAACTCTGTTGTAATTTAGCAATCTAC	1502
QY	6312	ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCTCCACACAATAGTCTGCT	6371
Db	1503	ATACACAACCAACTCCCAATGCATCAGGTGGGTAACTCTCCACACAATAGTCTGCT	1562
QY	6372	ACCCTCAGGAATATTTTGTCTGTGTGCTCCTCAGCCTATCGTTGTTGAATGGCTCTTC	6431
Db			

Db	1563	ACCCTCAGGAATATTTTGTCTGTGTGTACCTCAGCCTATCGTTGTTGAATGGCTCTTC	1622
QY	6432	AGAATCTATGTGCTTCTCTCAATTTAGTGGCCCCCATATGRCCTCTACACTGACACA	6491
Db	1623	AGAATCTATGTGCTTCTCTCAATTTAGTGGCCCCCATATGRCCTCTACACTGACACA	1682
QY	6492	TTTATACAGTTATGTCTATCTAAGCCCGCAACAAAGAGTACCCATTTCTTCTTTTGT	6551
Db	1683	TTTATACAGTTATGTCTATCTAAGCCCGCAACAAAGAGTACCCATTTCTTCTTTTGT	1742
QY	6552	TATAGGAGCAGGAGTGTAGGTGCACTAGTACTGGCATTCGGCGTATCAACAACCTCTAC	6611
Db	1743	TATAGGAGCAGGAGTGTAGGTGCACTAGTACTGGCATTCGGCGTATCAACAACCTCTAC	1802
QY	6612	TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTGCGCGACTC	6671
Db	1803	TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTGCGCGACTC	1862
QY	6672	CCTGCTCACTTGCAGAGTCAACTTAACTCCCTAGCAGCAGTAGTCTTCAAATTCGAAG	6731
Db	1863	CCTGCTCACTTGCAGAGTCAACTTAACTCCCTAGCAGCAGTAGTCTTCAAATTCGAAG	1922
QY	6732	AGCTTTAGACTTGTAAACCGCTGARAGAGGGGAACTGTTTATTTTATAGGGGAAGAATG	6791
Db	1923	AGCTTTAGACTTGTAAACCGCTGARAGAGGGGAACTGTTTATTTTATAGGGGAAGAATG	1982
QY	6792	CTGTTATTTATGTTTAAATCCGAATCGTCACTGAGAGAACTTAAAGAAATTCGAGATCG	6851
Db	1983	CTGTTATTTATGTTTAAATCCGAATCGTCACTGAGAGAACTTAAAGAAATTCGAGATCG	2042
QY	6852	AATCAACGCTAKAGCAGAGAGCTTCGAAACACCTGGACCCTGGGGCCCTCTCAGCCATG	6911
Db	2043	AATCAACGCTAKAGCAGAGAGCTTCGAAACACCTGGACCCTGGGGCCCTCTCAGCCATG	2102
QY	6912	GATGCCCTGGATTTCTCCCTTTTAGGACCTCTAGCAGCTATAATPATTTGCTCTCTCT	6971
Db	2103	GATGCCCTGGATTTCTCCCTTTTAGGACCTCTAGCAGCTATAATPATTTGCTCTCTCT	2162
QY	6972	TGGACCTGATCTTTTACCTCTTGTAACTTTGTCTCTTCCAGAAATCGAAGCTGTAA	7031
Db	2163	TGGACCTGATCTTTTACCTCTTGTAACTTTGTCTCTTCCAGAAATCGAAGCTGTAA	2222
QY	7032	ACTCAAAATGAGCCCAAGATGCGAGTCAAGACTTAAAGACTTACCGCAGACCCCTGACCG	7091
Db	2223	ACTCAAAATGAGCCCAAGATGCGAGTCAAGACTTAAAGACTTACCGCAGACCCCTGACCG	2282
QY	7092	GCCTGYTAGCCCAAGATCTGATGTTTAAATGACATCAAGGCAACCCCTCTGAGGAAATCTC	7151
Db	2283	GCCTGYTAGCCCAAGATCTGATGTTTAAATGACATCAAGGCAACCCCTCTGAGGAAATCTC	2342
QY	7152	AGCTGCACAACTCTACTACGCCCAATTCAGCAGGAAAGCAGTTAGAGGGTGTGCGGC	7211
Db	2343	AGCTGCACAACTCTACTACGCCCAATTCAGCAGGAAAGCAGTTAGAGGGTGTGCGGC	2402
QY	7212	AACCTCCCAACAGCAGCTTAGGTTTCTGTTGAGATGGGACCTGAGAGCAGGACTAG	7271
Db	2403	AACCTCCCAACAGCAGCTTAGGTTTCTGTTGAGATGGGACCTGAGAGCAGGACTAG	2462
QY	7272	CTGGATTTCTTAGGCTGATTAAGAAATCCYTAAGCCTAGSTGGGAGGTGACACATCCAC	7331
Db	2463	CTGGATTTCTTAGGCTGATTAAGAAATCCYTAAGCCTAGSTGGGAGGTGACACATCCAC	2522
QY	7332	CTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATTCAGAGAGCTCACTAAAT	7391
Db	2523	CTTTAAACACGGGGCTTGCAACTTAGYTCACACCTGACCAATTCAGAGAGCTCACTAAAT	2582
QY	7392	GCTAATTAGGCAAGACAGAGGAGTAAAGAAATAGCCAATCATYATTGCTGAGAGCACA	7451
Db	2583	GCTAATTAGGCAAGACAGAGGAGTAAAGAAATAGCCAATCATYATTGCTGAGAGCACA	2642
QY	7452	GCAGGAGGCAATGATCGGGATATAAACCCCAAGTTCGAGCCCGGCAACCGCAACCCCC	7511
Db	2643	GCAGGAGGCAATGATCGGGATATAAACCCCAAGTTCGAGCCCGGCAACCGCAACCCCC	2702



```

QY 7512 TTGGGTCCTCCCTTGTATGGAGCTCTGTTTCACTGCTATTT 7557
Db 7203 TTGGGTCCTCCCTTGTATGGAGCTCTGTTTCACTGCTATTT 2748

RESULT 2
LOCUS CR617248 2749 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0D1022YJ18 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR617248
VERSION 1
KEYWORDS HTc; CnSLT cDNA
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 2749)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Paraday Avenue
2 (bases 1 to 2749)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
location/Qualifiers
1..2749
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1022YJ18"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 29.1%; Score 2176.4; DB 6; Length 2749;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 2167; Conservative 34; Mismatches 0; Indels 1; Gaps 1;

QY 5352 GATACATCTGGGAGGACCTACCCAGTCATTTATYACCCCACTGGGTTAAAGTG 5411
Db 549 GATACATCTGGGAGGACCTACCCAGTCATTTATYACCCCACTGGGTTAAAGTG 608

QY 5412 GCTGGAGTGGAGTCTTGGGATACATCACTGAGTCAAAATCTGGGATCTGCCAAGGAA 5471
Db 609 GCTGGAGTGGAGTCTTGGGATACATCACTGAGTCAAAATCTGGGATCTGCCAAGGAA 668

QY 5472 CCTGAAATCCAGGAGACACGCTAGCTATTCCTGGAACCTCTAGAGATTGGCCCTG 5531
Db 669 CCTGAAATCCAGGAGACACGCTAGCTATTCCTGGAACCTCTAGAGATTGGCCCTG 728

QY 5532 CTCCTCAAAACAACACGAGGAGAAAGTAACATAAATCAATATCCCCATGSCCTCCC 5591
Db 729 CTCCTCAAAACAACACGAGGAGAAAGTAACATAAATCAATATCCCCATGSCCTCCC 787

QY 5592 TTATCATATTTTCTCTKTAAGTSTTTTAAACCTSTTTCACTCTCACTGCAACCCCTCC 5651
Db 788 TTATCATATTTTCTCTKTAAGTSTTTTAAACCTSTTTCACTCTCACTGCAACCCCTCC 847

QY 5652 ATGCCGCTGTATGACAGTAGCTCCCTVACCMAGAGTTTCTATGGAGATGAGCGTCC 5711
Db 5652 ATGCCGCTGTATGACAGTAGCTCCCTVACCMAGAGTTTCTATGGAGATGAGCGTCC 5711

```

```

Db 848 ATGCCGCTGTATGACAGTAGCTCCCTTACCAAGAGTTTCTATGGAGATGAGCGTCC 907
QY 5712 CGGAAATATTGATGCCCAATCGTATAGGAGTCTTTSTAAGGGAACCCACCTTCACTGC 5771
Db 908 CGGAAATATTGATGCCCAATCGTATAGGAGTCTTTSTAAGGGAACCCACCTTCACTGC 967
QY 5772 CCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTTGATGATGCAATATAC 5831
Db 968 CCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTTGATGATGCAATATAC 1027
QY 5832 TCATATTTCGACAGGAAAAATGATTAATCTAGTTCTCTGGAGGACTTTGGAGTCACTGT 5891
Db 1028 TCATATTTCGACAGGAAAAATGATTAATCTAGTTCTCTGGAGGACTTTGGAGTCACTGT 1087
QY 5892 CTGTTGGACTTACTTCA CCCCACCTGTTATGTTGATGGGGTGGAGTTCAAGATCAGGC 5951
Db 1088 CTGTTGGACTTACTTCA CCCCACCTGTTATGTTGATGGGGTGGAGTTCAAGATCAGGC 1147
QY 5952 AAGAGAAAAACATGTAAAAAGAGTAATCTCCCACTCACCCGGGTACATGGCACTCTAG 6011
Db 1148 AAGAGAAAAACATGTAAAAAGAGTAATCTCCCACTCACCCGGGTACATGGCACTCTAG 1207
QY 6012 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTAACCCATCTCGCCT 6071
Db 1208 CCCCTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTAACCCATCTCGCCT 1267
QY 6072 GGTAAAGCCTATTTAATACACCCCTCACTGGGGTCCATGAGGTCTCGGCCCAAAACCTTAC 6131
Db 1268 GGTAAAGCCTATTTAATACACCCCTCACTGGGGTCCATGAGGTCTCGGCCCAAAACCTTAC 1327
QY 6132 TAACTGTTGGATATGCTCCCTGAACTTTCARGGCAATATGTTCAATCCCTGTACTGA 6191
Db 1328 TAACTGTTGGATATGCTCCCTGAACTTTCARGGCAATATGTTCAATCCCTGTACTGA 1387
QY 6192 ACAATGGAAACAATTCAGACAGAAATAAACAACCACTTCCTGTTAGTAGGACCTCTTGT 6251
Db 1388 ACAATGGAAACAATTCAGACAGAAATAAACAACCACTTCCTGTTAGTAGGACCTCTTGT 1447
QY 6252 TTCCAATSTGGAAATAAACCCATACCTCAAACTCACTGTGTAAATTTAGCAATACTAC 6311
Db 1448 TTCCAATCTGGAAATAAACCCATACCTCAAACTCACTGTGTAAATTTAGCAATACTAC 1507
QY 6312 ATACAACAACCACTCCCAATGATCAGGTGGGTAACTCTCTCCAACAATAAGTCTGCT 6371
Db 1508 ATACAACAACCACTCCCAATGATCAGGTGGGTAACTCTCTCCAACAATAAGTCTGCT 1567
QY 6372 ACCCTCAGGAATATTTTGTCTGTGTGTTACCTCAGCCTATCTGTTGAAATGGCTCTTC 6431
Db 1568 ACCCTCAGGAATATTTTGTCTGTGTGTTACCTCAGCCTATCTGTTGAAATGGCTCTTC 1627
QY 6432 AGAATCTATGTCTTCTCTCTCACTTCTAGTGCCCGCCYATGRCCTATCTACACTGAACA 6491
Db 1628 AGAATCTATGTCTTCTCTCTCACTTCTAGTGCCCGCCYATGRCCTATCTACACTGAACA 1687
QY 6492 TTTATACAGTTATGTATATCTAAGCCCGCAACAAAAGAGTACCCTTCTTCTTTTGT 6551
Db 1688 TTTATACAGTTATGTATATCTAAGCCCGCAACAAAAGAGTACCCTTCTTCTTTTGT 1747
QY 6552 TATAGAGCAGAGTGTCTAGGTGCACTAGGTGCTGGCAATGGCGGTATCACAACCTCTAC 6611
Db 1748 TATAGAGCAGAGTGTCTAGGTGCACTAGGTGCTGGCAATGGCGGTATCACAACCTCTAC 1807
QY 6612 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGACATGGAACGGGTGCGGACTC 6671
Db 1808 TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGACATGGAACGGGTGCGGACTC 1867
QY 6672 CCTGCTCACCTTGCAGAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTTCRAAATCGAAG 6731
Db 1868 CCTGCTCACCTTGCAGAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTTCRAAATCGAAG 1927
QY 6732 AGCTTTAGACTVGTAAACCGCTGAPAGAGGGGAAACCTGTTTATTTTATAGGGGAAATG 6791
Db 1928 AGCTTTAGACTVGTAAACCGCTGAPAGAGGGGAAACCTGTTTATTTTATAGGGGAAATG 1987

```

```
QY 6792 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAATTCAGATCG 6851
Db 1988 CTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAATTCAGATCG 2047
QY 6852 AATCAACGCTAKAGCAGAGAGCTTCGAAACACTGAGACCTGGGCGCTCTCAGCCCATG 6911
Db 2048 AATCAACGCTAGAGCAGAGAGCTTCGAAACACTGAGACCTGGGCGCTCTCAGCCCATG 2107
QY 6912 GATGCCCTGGATTCCTCCCTCTTATAGACCTCTAGCAGCTATAATATTGCTACTCTCTT 6971
Db 2108 GATGCCCTGGATTCCTCCCTCTTATAGACCTCTAGCAGCTATAATATTGCTACTCTCTT 2167
QY 6972 TGGACCTCTGATCTTTTACCTCTCTTTAACTTTCTCTTCCAGAAATCGAGCTGTA 7031
Db 2168 TGGACCTCTGATCTTTTAACTCTCTTTAACTTTCTCTTCCAGAAATCGAGCTGTA 2227
QY 7032 ACTCAAAATGGAGCCCAAGATGACAGTCCAAAGACTAAGATCTACCGCAGACCCCTGAGCG 7091
Db 2228 ACTCAAAATGGAGCCCAAGATGACAGTCCAAAGACTAAGATCTACCGCAGACCCCTGAGCG 2287
QY 7092 GCGTGYTAGCCCAAGATCTGATGTTTAATGACATCAAAAGCACCCCTCTCTAGGAAATCTC 7151
Db 2288 GCGTGTAGCCCAAGATCTGATGTTTAATGACATCAAAAGCACCCCTCTCTAGGAAATCTC 2347
QY 7152 AGCTGCACAACTCTACTAGCCGCCAATTCAGCAGAGAGCACTTACGAGCGGTGTCGCGC 7211
Db 2348 AGCTGCACAACTCTACTAGCCGCCAATTCAGCAGAGAGCACTTACGAGCGGTGTCGCGC 2407
QY 7212 RACCTCCCAACAGCAGCTTAGGTTTCTCTGTTGAGATGGGAGCTGAGAGCAGGACTAG 7271
Db 2408 AACCTCCCAACAGCAGCTTAGGTTTCTCTGTTGAGATGGGAGCTGAGAGCAGGACTAG 2467
QY 7272 CTGGATTTCTTAGGCTGATYTAAGAAATCCYTAAGCCTTAGSTGGAAAGGTGACCAATCCAC 7331
Db 2468 CTGGATTTCTTAGGCTGATYTAAGAAATCCYTAAGCCTTAGSTGGAAAGGTGACCAATCCAC 2527
QY 7332 CTTTAAACACGGGGCTTGCACTTAGTCTACACCTGACCAATCAGAGAGCTCACTAAAT 7391
Db 2528 CTTTAAACACGGGGCTTGCACTTAGTCTACACCTGACCAATCAGAGAGCTCACTAAAT 2587
QY 7392 GCTAATTAGGCAAGAGAGAGTAAAGAAATAGCCAAATCATYTTATGCTGAGAGCACA 7451
Db 2588 GCTAATTAGGCAAGAGAGAGTAAAGAAATAGCCAAATCATYTTATGCTGAGAGCACA 2647
QY 7452 GCAGGAGGACAATGATCGGGATATAAACCCAAAGTTCGAGCCCGGCAACGGCAACCC 7511
Db 2648 GCAGGAGGACAATGATCGGGATATAAACCCAAAGTTCGAGCCCGGCAACGGCAACCC 2707
QY 7512 TTTGGTCCCTCCCTTTGTTATGGGAGCTCTGTTTCATGCT 7553
Db 2708 TTTGGTCCCTCCCTTTGTTATGGGAGCTCTGTTTCATGCT 2749
```

## RESULT 3

```
CR622175 2500 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DI051YM13 of Placenta Cot 25-normalized
DEFINITION full-length cDNA clone CS0DI051YM13 of Placenta Cot 25-normalized
ACCESSION CR622175
VERSION CR622175.1 GI:50502982
KEYWORDS HTC; CNSLT_CDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2500)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
```

```
Paraday Avenue
2 (bases 1 to 2500)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
1..2500
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI051YM13"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 29.0%; Score 2174.4; DB 6; Length 2500;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 2165; Conservative 34; Mismatches 0; Indels 1; Gaps 1;
QY 5352 GATACATCCTCGGAAGGACCCCTACCCAGTCATTTTATYACCCCAACTGCGGTTAAAGTG 5411
Db 302 GATACATCCTCGGAAGGACCCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTG 361
QY 5412 GCTGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAATCTTGATACTGCCAAGGAA 5471
Db 362 GCTGAGTGGAGTCTTGGATACATCACACTTGAGTCAAAATCTTGATACTGCCAAGGAA 421
QY 5472 CTTGAAATCAGAGAGCAACGCTAGCTATTCTTGTGAACCTCTAGAGAGATTGGCGCTG 5531
Db 422 CTTGAAATCAGAGAGCAACGCTAGCTATTCTTGTGAACCTCTAGAGAGATTGGCGCTG 481
QY 5532 CTCTTCAAAACAACAGGAGGAGTAACCTAAATCATAAATCAAAATCAAAATCAAAATCAAAAT 5591
Db 482 CTCTTCAAAACAACAGGAGGAGTAACCTAAATCATAAATCAAAATCAAAATCAAAATCAAAAT 540
QY 5592 TTATCATATTTTCTCTKTASTGTTSTTTTACCCTTCTCTCACTCTCACTGCAACCCCTCC 5651
Db 541 TTATCATATTTTCTCTTACTGTTCTTTTACCCCTCTTCTCACTCTCACTGCAACCCCTCC 600
QY 5652 ATGCGCTGTATGACAGTAGCTCCCTTACCMAGAGTTTCTATGGAGAGATGACAGCTCC 5711
Db 601 ATGCGCTGTATGACAGTAGCTCCCTTACCMAGAGTTTCTATGGAGAGATGACAGCTCC 660
QY 5712 CGGAATATTGATGCCCATCTGATAGAGTCTTTSTAAGGGAACCCCACTTCACTGC 5771
Db 661 CGGAATATTGATGCCCATCTGATAGAGTCTTTCTAAGGGAACCCCACTTCACTGC 720
QY 5772 CCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGATGATGCAATAC 5831
Db 721 CCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGATGATGCAATAC 780
QY 5832 TCATTATTGGACAGGAAATAATGATTAAATCTCTAGTTGTCTTGAGAGATTTGGAGTCACTGT 5891
Db 781 TCATTATTGGACAGGAAATAATGATTAAATCTCTAGTTGTCTTGAGAGATTTGGAGTCACTGT 840
QY 5892 CTGTTGGACTTTACTTACCCCAACTGTTATGCTGATGGGGGTGAGTTCAAGATCAGGC 5951
Db 841 CTGTTGGACTTTACTTACCCCAACTGTTATGCTGATGGGGGTGAGTTCAAGATCAGGC 900
QY 5952 AAGAGAAAAACATGTAAAGAGTAATCTCCCAACTCACCCGGGTACATGCACTCTAG 6011
Db 901 AAGAGAAAAACATGTAAAGAGTAATCTCCCAACTCACCCGGGTACATGCACTCTAG 960
QY 6012 CCCCTCAAGAGGACTAGATCTCTCAAAACTACATCAAAACCCCTCCGATACCCATCTGCCT 6071
Db 961 CCCCTCAAGAGGACTAGATCTCTCAAAACTACATCAAAACCCCTCCGATACCCATCTGCCT 1020
```

```
QY 6072 GGTAAAGCCTATTAAATACCAACCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCCTAC 6131
Db 1021 GGTAAAGCCTATTAAATACCAACCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCCTAC 1080
QY 6132 TAACTGTTGGATATGCTCCCTCCCTGAACTTCAAGCCCATATGTTTCAATCCCTGTAAGTGA 6191
Db 1081 TAACTGTTGGATATGCTCCCTCCCTGAACTTCAAGCCCATATGTTTCAATCCCTGTAAGTGA 1140
QY 6192 ACAATGGAACTCTCAGCAAGAAATAAACAACACTTCGTTTAAAGTAGGAGCTCTTGT 6251
Db 1141 ACAATGGAACTCTCAGCAAGAAATAAACAACACTTCGTTTAAAGTAGGAGCTCTTGT 1200
QY 6252 TTCCAATSTGGAAATACCAATACCTCAACCTCACTGTGTAAATTTAGCAATTAATCTTGT 6311
Db 1201 TTCCAATSTGGAAATACCAATACCTCAACCTCACTGTGTAAATTTAGCAATTAATCTTGT 1260
QY 6312 ATACACAACCAACTCCCAATGCAATCAGGTGGGTAACTCCCTCCACACAATAAGTCTGCGCT 6371
Db 1261 ATACACAACCAACTCCCAATGCAATCAGGTGGGTAACTCCCTCCACACAATAAGTCTGCGCT 1320
QY 6372 ACCCTCAGGAATATTTTGTCTGTGTAACCTCAGCCTATCGTTGTTGAATGGCTCTTC 6431
Db 1321 ACCCTCAGGAATATTTTGTCTGTGTAACCTCAGCCTATCGTTGTTGAATGGCTCTTC 1380
QY 6432 AGAATCTATGTGCTCTCTCATTTAGTGCCCCCCTATGRCATCTACACTGAAACAAGA 6491
Db 1381 AGAATCTATGTGCTCTCTCATTTAGTGCCCCCCTATGRCATCTACACTGAAACAAGA 1440
QY 6492 TTTTATACAGTTATGTCATATCTAAGCCCGCAACAAAGAGTACCCATCTCTCTTTTGT 6551
Db 1441 TTTTATACAGTTATGTCATATCTAAGCCCGCAACAAAGAGTACCCATCTCTCTTTTGT 1500
QY 6552 TATAGGAGCAGGAGTCTAGGTGCACTAGGTACTGGCATTTGGCGGTATCAACACTCTAC 6611
Db 1501 TATAGGAGCAGGAGTCTAGGTGCACTAGGTACTGGCATTTGGCGGTATCAACACTCTAC 1560
QY 6612 TCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGGTGCGCACTC 6671
Db 1561 TCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGGTGCGCACTC 1620
QY 6672 CTTGGTCACTTGCAGAGTCACTTAACCTCCCTAGCAGCAGTAGTCTTCRAAATCGAAG 6731
Db 1621 CTTGGTCACTTGCAGAGTCACTTAACCTCCCTAGCAGCAGTAGTCTTCRAAATCGAAG 1680
QY 6732 AGCTTTAGACTTGTCAACCGCTGAGAGAGGGGAACCTGTTTATTTTATAGGGGAAGAATG 6791
Db 1681 AGCTTTAGACTTGTCAACCGCTGAGAGAGGGGAACCTGTTTATTTTATAGGGGAAGAATG 1740
QY 6792 CTGTTATTAATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAATTCAGATCG 6851
Db 1741 CTGTTATTAATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAATTCAGATCG 1800
QY 6852 AATACAACTAGAGAGAGAGTCTGAAACACTGACACCTGGGGCTCTCAGCCRAATG 6911
Db 1801 AATACAACTAGAGAGAGAGTCTGAAACACTGACACCTGGGGCTCTCAGCCRAATG 1860
QY 6912 GATGCCCTGATTTCTCCCTTCTTAGGACTCTAGCAGCTATAATTTCTACTCTCTCTT 6971
Db 1861 GATGCCCTGATTTCTCCCTTCTTAGGACTCTAGCAGCTATAATTTCTACTCTCTCTT 1920
QY 6972 TGGACCTGTATCTTTTACCTCTCTGTTAACTTTGTTCTTCCAGAAATCGAAGCTGTAA 7031
Db 1921 TGGACCTGTATCTTTTACCTCTCTGTTAACTTTGTTCTTCCAGAAATCGAAGCTGTAA 1980
QY 7032 ACTACAAATGGAGCCCAAGATGAGTCCAGACTAAGACTAATACCCAGACCCCTGAGACG 7091
Db 1981 ACTACAAATGGAGCCCAAGATGAGTCCAGACTAAGACTAATACCCAGACCCCTGAGACG 2040
QY 7092 GCTGTGTAGCCCAAGATCTGATGTTAATGACATCAAGGACCCCTCTCTGAGAAATCTC 7151
Db 2041 GCTGTGTAGCCCAAGATCTGATGTTAATGACATCAAGGACCCCTCTCTGAGAAATCTC 2100
```

```
QY 7152 AGTGTCACAACTCTACTAGCCGCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTGGCC 7211
Db 2101 AGTGTCACAACTCTACTAGCCGCCAATTCAGCAGGAAGCAGTTAGAGCGGTSGTGGCC 2160
QY 7212 AACCTCCCAACAGCACTTAGGTTTCTGTTGAGATGGGACTCAGAGACAGGACTAG 7271
Db 2161 AACCTCCCAACAGCACTTAGGTTTCTGTTGAGATGGGACTCAGAGACAGGACTAG 2220
QY 7272 CTGGAATTTCTAGGCTGATTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCAATCCAC 7331
Db 2221 CTGGAATTTCTAGGCTGATTAAGAATCCYTAAGCCTAGSTGGGAAGGTGACCAATCCAC 2280
QY 7332 CTTTAAACAGCGGGCTTGCAACTTAGTTCACACTGACCAATCAGAGAGCTCACTAAAT 7391
Db 2281 CTTTAAACAGCGGGCTTGCAACTTAGTTCACACTGACCAATCAGAGAGCTCACTAAAT 2340
QY 7392 GCTTAATTAGCAAGACAGGAGTAAAGAAATAGCCAAATCATVTTATTTGWTGAGACACA 7451
Db 2341 GCTTAATTAGCAAGACAGGAGTAAAGAAATAGCCAAATCATCTATTGCTGAGAGACACA 2400
QY 7452 GCAGGAGGACAAATGATCGGGATATAAACCCAAAGTTCGAGCCGCAACGGCAACCCCC 7511
Db 2401 GCAGGAGGACAAATGATCGGGATATAAACCCAAAGTTCGAGCCGCAACGGCAACCCCC 2460
QY 7512 TTTGGTCCCTCCCTCTTGTATGGGAGCTCTGTTTTCATG 7551
Db 2461 TTTGGTCCCTCCCTCTTGTATGGGAGCTCTGTTTTCATG 2500
```

## RESULT 4

CR625046

LOCUS

DEFINITION

2748 bp mRNA linear HTC 21-JUL-2004

full-length cDNA clone CS0D1044YK06 of Placenta Cot 25-normalized

of Homo sapiens (human)

ACCESSION

CR625046

VERSION

CR625046.1 GI:50505853

KEYWORDS

HTC; CNSLT\_cDNA.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE

1. (bases 1 to 2748)

AUTHORS

Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished

REMARK

Contact : Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue

Genoscope.

REFERENCE

2. (bases 1 to 2748)

AUTHORS

Direct Submission

JOURNAL

Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

COMMENT

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen.

## FEATURES

source

1. 2748

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0D1044YK06"

/tissue type="Placenta Cot 25-normalized"

/plasmid="pCMVSPORT\_6"

ORIGIN

Query Match

Best Local Similarity

Matches 2160; Conservative

29.0%; Score 2169.4; DB 6; Length 2748;

98.4%; Pred. No. 0;

34; Mismatches

0; Indels 1; Gaps 1;

QY	5352	GATACATCCTGGGAGGACCCCTACCCAGCTCATTTTATYTA	CCCCAACTGCGGTTAAAGTG	5411
Db	555	GATACATCCTGGGAGGACCCCTACCCAGCTCATTTTAT	TACCCCACTGCGGTTAAAGTG	614
QY	5412	GCTGGAGTGGAGTCTTTGGATACATACACTTGAGTCAAA	TCTGGATACTGCAAAAGGAA	5471
Db	615	GCTGGAGTGGAGTCTTTGGATACATACACTTGAGTCAAA	TCTGGATACTGCAAAAGGAA	674
QY	5472	CCTGAAATCCAGGAGACAACGCTAGCTATTCTCTGTAAC	CTCTAGAGATTTGCGCCTG	5531
Db	675	CCTGAAATCCAGGAGACAACGCTAGCTATTCTCTGTAAC	CTCTAGAGATTTGCGCCTG	734
QY	5532	CTCTTCAACAACAACGAGGAGAAAGTAACTAAATCATAA	ATCCCCATGGSSCTCC	5591
Db	735	CTCTTCAACAACAACGAGGAGAAAGTAACTAAATCATAA	ATCCCCATGGSSCTCC	793
QY	5592	TTATCATATTTTCTCTKTASTGTTSTTTTACCTSTTTTCA	CTCTCACTGCAACCCCTCC	5651
Db	794	TTATCATATTTTCTCTKTASTGTTSTTTTACCTSTTTTCA	CTCTCACTGCAACCCCTCC	853
QY	5652	ATGCGCGTGTATGACCAAGTAGCTCCCTYACCMAGAGTTT	CTATGGAGAAATGCAGCGTCC	5711
Db	854	ATGCGCGTGTATGACCAAGTAGCTCCCTYACCMAGAGTTT	CTATGGAGAAATGCAGCGTCC	913
QY	5712	CGGAAATATGTAGTCCCATGCTATAGGAGTCTTTSTAAGG	AAACCCCACTTCACTGC	5771
Db	914	CGGAAATATGTAGTCCCATGCTATAGGAGTCTTTSTAAGG	AAACCCCACTTCACTGC	973
QY	5772	CCACACCCATATGCCCCGCAACGCTATCACTCTGCCACT	CTTTTCATGCAATGCATGCAAAATAC	5831
Db	974	CCACACCCATATGCCCCGCAACGCTATCACTCTGCCACT	CTTTTCATGCAATGCATGCAAAATAC	1033
QY	5832	TCATTTATGGGACAGGAAATGATTAATCTAGTTGTCTGCT	GAGGACTTGGAGTCACTGT	5891
Db	1034	TCATTTATGGGACAGGAAATGATTAATCTAGTTGTCTGCT	GAGGACTTGGAGTCACTGT	1093
QY	5892	CTGTTGGACTTACTTCAACCAACTGGTATGTCTGATGGG	GGTGGAGTTCAAGATCAGGC	5951
Db	1094	CTGTTGGACTTACTTCAACCAACTGGTATGTCTGATGGG	GGTGGAGTTCAAGATCAGGC	1153
QY	5952	AAAGAAACATGTAAGAGAGTAACTCCCAACTCACCSGG	TACATGGCACCTCTAG	6011
Db	1154	AAAGAAACATGTAAGAGAGTAACTCCCAACTCACCSGG	TACATGGCACCTCTAG	1213
QY	6012	CCCTCTACAAAGGACTAGATCTCTCAAAACTACATGAAA	CCCTCCGTACCCATACCTCGCT	6071
Db	1214	CCCTCTACAAAGGACTAGATCTCTCAAAACTACATGAAA	CCCTCCGTACCCATACCTCGCT	1273
QY	6072	GGTAAGCCTATTTAATACACCTCACTGGGCTTCAATAGG	TCTCGGCCCAAAACCCCTAC	6131
Db	1274	GGTAAGCCTATTTAATACACCTCACTGGGCTTCAATAGG	TCTCGGCCCAAAACCCCTAC	1333
QY	6132	TAACTGTTGGATATGCCTCCCTGAACTTTCAGGCCATAT	GTTCATCCCTGTACCTGA	6191
Db	1334	TAACTGTTGGATATGCCTCCCTGAACTTTCAGGCCATAT	GTTCATCCCTGTACCTGA	1393
QY	6192	ACAAATGGAAACAATTCAGCACAGAAATAAACACCACT	TCCGCTTTTATGAGCACTCTTGT	6251
Db	1394	ACAAATGGAAACAATTCAGCACAGAAATAAACACCACT	TCCGCTTTTATGAGCACTCTTGT	1453
QY	6252	TTCCAAATSTGGAATAACCACTACCTCAAACTCACTGTA	AGATTTTGAATTAATCTAC	6311
Db	1454	TTCCAAATSTGGAATAACCACTACCTCAAACTCACTGTA	AGATTTTGAATTAATCTAC	1513
QY	6312	ATACACCAACCACTCCCAATGATCAGGTGGGTAACTCT	CCCAACCAATAGTCTGCT	6371
Db	1514	ATACACCAACCACTCCCAATGATCAGGTGGGTAACTCT	CCCAACCAATAGTCTGCT	1573
QY	6372	ACCCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTAT	CGTTGTTGAATGGCTCTTC	6431
Db	1574	ACCCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTAT	CGTTGTTGAATGGCTCTTC	1633

QY	6432	AGAATCTATGTGCTTCTCTCAITCTTAGTGCCCCCVATGR	CCCTATCTACACTGAACAAGA	6491
Db	1634	AGAATCTATGTGCTTCTCTCAITCTTAGTGCCCCCVATGR	CCCTATCTACACTGAACAAGA	1693
QY	6492	TTTATACAGTTTATGTCTATCTAAGCCCGCAACAAAGAG	TACCCATCTTCTCTTTTGT	6551
Db	1694	TTTATACAGTTTATGTCTATCTAAGCCCGCAACAAAGAG	TACCCATCTTCTCTTTTGT	1753
QY	6552	TATAGGAGCAGGAGTGTCTAGTGCACTAGGTACTGGCA	TGTGGCGGTATCAAACTCTAC	6611
Db	1754	TATAGGAGCAGGAGTGTCTAGTGCACTAGGTACTGGCA	TGTGGCGGTATCAAACTCTAC	1813
QY	6612	TCAGTTCTACTACAAAATCTATCTCAAGAACTAAATGG	GGGACATGAAACGGGTCCGCACTC	6671
Db	1814	TCAGTTCTACTACAAAATCTATCTCAAGAACTAAATGG	GGGACATGAAACGGGTCCGCACTC	1873
QY	6672	CCTGGTCAACCTTGCAGAGATCAACTTAACTCCTTAGC	AGCAGTAGTCTTTCRAAATCGAAG	6731
Db	1874	CCTGGTCAACCTTGCAGAGATCAACTTAACTCCTTAGC	AGCAGTAGTCTTTCRAAATCGAAG	1933
QY	6732	AGCTTTAGACTYTGCTAAACCGCTGAGAGAGGGGAACT	GTGTTTATTTTTAGGGGAAAGATG	6791
Db	1934	AGCTTTAGACTYTGCTAAACCGCTGAGAGAGGGGAACT	GTGTTTATTTTTAGGGGAAAGATG	1993
QY	6792	CTGTATTATGTAAATCAATCCGGAATCGTCACTGAGAA	AGCTTAAAGAAATTCGAGATCG	6851
Db	1994	CTGTATTATGTAAATCAATCCGGAATCGTCACTGAGAA	AGCTTAAAGAAATTCGAGATCG	2053
QY	6852	AATCAACGTAAGACAGAGAGCTTTCGAAAACACTTGG	ACCTCTGGGCGCTCTCAGCCRAATG	6911
Db	2054	AATCAACGTAAGACAGAGAGCTTTCGAAAACACTTGG	ACCTCTGGGCGCTCTCAGCCRAATG	2113
QY	6912	GATGCCCTGGATCTCCCTCTTATAGGACCTCTAGCAGC	TATATATTTGCTACTCTCTCTT	6971
Db	2114	GATGCCCTGGATCTCCCTCTTATAGGACCTCTAGCAGC	TATATATTTGCTACTCTCTCTT	2173
QY	6972	TGGACCTGTATCTTTTACCTCTCTTGTAACTTTTAA	CTTTCTCTTCCAGAAATCGAAGCTGTAA	7031
Db	2174	TGGACCTGTATCTTTTACCTCTCTTGTAACTTTTAA	CTTTCTCTTCCAGAAATCGAAGCTGTAA	2233
QY	7032	ACTCAAAATGAGGCCCAAGATGCAAGTCAAGACTTAA	AGACTCTACCGCAGACCCCTCGACCG	7091
Db	2234	ACTCAAAATGAGGCCCAAGATGCAAGTCAAGACTTAA	AGACTCTACCGCAGACCCCTCGACCG	2293
QY	7092	GCCTGTAGCCCAAGATCTGATGTTAATGACATCAAA	AGGGAACCCCTCTCGAGGAAATCTC	7151
Db	2294	GCCTGTAGCCCAAGATCTGATGTTAATGACATCAAA	AGGGAACCCCTCTCGAGGAAATCTC	2353
QY	7152	AGCTGCAAACTCTACTACGCCCCCAATTCAGCAGGA	AGCAGTTAGAGGGTSGTCCGCC	7211
Db	2354	AGCTGCAAACTCTACTACGCCCCCAATTCAGCAGGA	AGCAGTTAGAGGGTSGTCCGCC	2413
QY	7212	AACCTCCCAACAGCAGCTTAGTCTTCTGTTAGATGG	GGAATCTGAGAGACAGGACTAG	7271
Db	2414	AACCTCCCAACAGCAGCTTAGTCTTCTGTTAGATGG	GGAATCTGAGAGACAGGACTAG	2473
QY	7272	CTGGATTTCTTAGGCTGAYTAAGAAATCCYTAAGCCT	TAGSTGGGAAAGGTGACACATCCAC	7331
Db	2474	CTGGATTTCTTAGGCTGAYTAAGAAATCCYTAAGCCT	TAGSTGGGAAAGGTGACACATCCAC	2533
QY	7332	CTTTTAAACAACCGGGCTTGCACCTTAGTTCACACCT	TGACCAATCAGAGAGCTCACTAAAT	7391
Db	2534	CTTTTAAACAACCGGGCTTGCACCTTAGTTCACACCT	TGACCAATCAGAGAGCTCACTAAAT	2593
QY	7392	GCTAATTAGGCAAAAGACAGGAGGTAAGAAATAGCCA	ATCATYTTATTTGWTGAGAGCA	7451
Db	2594	GCTAATTAGGCAAAAGACAGGAGGTAAGAAATAGCCA	ATCATYTTATTTGWTGAGAGCA	2653
QY	7452	GCAGGAGGAAATATGATCGGATATPAAACCAAGT	YTTTCGAGCCCGGCAACCGCCCC	7511
Db	2654	GCAGGAGGAAATATGATCGGATATPAAACCAAGT	YTTTCGAGCCCGGCAACCGCCCC	2713
QY	7512	TTTGGGTCCTCCCTCTTGTATGGGAGCTCTGTTT	7546	

Db	2714	TTTTGGTCCCTCCCTTTGTATGGAGCTCTGT	2748
RESULT 5			
CR613169			
LOCUS			
DEFINITION	CR613169	2716 bp	linear HTC 21-JUL-2004
ACCESSION	full-length cDNA clone CS0DE013Y120 of Placenta of Homo sapiens (human)		
VERSION	CR613169.1	GI:50493976	
KEYWORDS	HTC; CNSLT cDNA		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 2716)		
TITLE	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.		
JOURNAL	Full-length cDNA libraries and normalization		
REMARK	Unpublished		
REFERENCE	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue		
AUTHORS	2 (bases 1 to 2716)		
TITLE	Genoscope.		
JOURNAL	Direct Submission		
COMMENT	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)		
FEATURES	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.		
source	Location/Qualifiers		
	1. .2716		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="CS0DE013Y120"		
	/tissue_type="Placenta"		
	/plasmid="pCMVSPORT_6"		
ORIGIN			
Query Match	28.7%;	Score 2147.4;	DB 6; Length 2716;
Best Local Similarity	98.4%;	Pred. No. 0;	
Matches 2138;	Conservative	34;	Mismatches 0; Indels 1; Gaps 1;
QY	5352	GATACATCTCGGGAAGGACCTCACCCAGTCA	TTTTATYATACCCCACTGCGGTAAAGTG 5411
Db	545	GATACATCTCTGGGAAGGACCTCACCCAGTCA	TTTTATCTACCCCAACTGCGGTAAAGTG 604
QY	5412	GCTGGAGTGAGTGCTTGGATACATCACATCTTGAGTCAAAATCCTGGATAC	TGCCCAAGGAA 5471
Db	605	GCTGGAGTGAGTGCTTGGATACATCACATCTTGAGTCAAAATCCTGGATAC	TGCCCAAGGAA 664
QY	5472	CCTGAAATCCAGAGACAAACGCTAGCTATTCCTGTGAACTCTAGAGGATTTGC	GCCTG 5531
Db	665	CCTGAAATCCAGAGACAAACGCTAGCTATTCCTGTGAACTCTAGAGGATTTGC	GCCTG 724
QY	5532	CTCTTCAACACAAACACAGAGAGAAAGTAAATCAAAATCCCAATGCCCTCC	5591
Db	725	CTCTTCAACACAAACACAGAGAGAAAGTAAATCAAAATCCCAATGCCCTCC	783
QY	5592	TTATCATATTTTCTCTKTASTGTSTTTTACCCCTSTTTTCACTCTCACTGCACCCCTCC	5651
Db	784	TTATCATATTTTCTCTKTACTGTCTTTTACCCCTCTTTTCACTCTCACTGCACCCCTCC	843
QY	5652	ATGCCGCTGTATGACCAAGTAGTCTCCCTYACCMAGAGTTTCTATGAGAAATGAGCGTCC	5711
Db	844	ATGCCGCTGTATGACCAAGTAGTCTCCCTTCAAGAGTTTCTATGAGAAATGAGCGTCC	903

Qy	5712	CGGAAATATTGATGCCCATCGTATAGAGTCTTTTSTAAGGGAACCCCACTTCACTGC	5771
Db	904	CGGAAATATTGATGCCCATCGTATAGAGTCTTTTSTAAGGGAACCCCACTTCACTGC	963
Qy	5772	CCACCCATATGCCCCGCAACTGCTATCACTCTGCCACCTTTTGGCATGCATGCAATAC	5831
Db	964	CCACCCATATGCCCCGCAACTGCTATCACTCTGCCACCTTTTGGCATGCATGCAATAC	1023
Qy	5832	TCATTATTGGACAGGAAAAATGATTAATTCCTAGTTGTCTTGGAGGACTTGGAGTCACTGT	5891
Db	1024	TCATTATTGGACAGGAAAAATGATTAATTCCTAGTTGTCTTGGAGGACTTGGAGTCACTGT	1083
Qy	5892	CTGTTGACTTACTTCAACCAACTGCTATGTCGTATAGTTGTCTTGGAGGACTTGGAGTCACTGT	5951
Db	1084	CTGTTGACTTACTTCAACCAACTGCTATGTCGTATAGTTGTCTTGGAGGACTTGGAGTCACTGT	1143
Qy	5952	AAGAGAAAAACATGTAAAGAAGTAATCTCCCAACTCACCGGGGTACATCGCACTCTAG	6011
Db	1144	AAGAGAAAAACATGTAAAGAAGTAATCTCCCAACTCACCGGGGTACATCGCACTCTAG	1203
Qy	6012	CCCCTACAAAGGACTAGATCTCTCAAAATACATGAACCCCTCCGTACCCATATCTGCCT	6071
Db	1204	CCCCTACAAAGGACTAGATCTCTCAAAATACATGAACCCCTCCGTACCCATATCTGCCT	1263
Qy	6072	GGTAAGCCTATTAAATACACACCTCACTGGGTCCATAGAGTCTCGGCCCAAAACCCCTAC	6131
Db	1264	GGTAAGCCTATTAAATACACACCTCACTGGGTCCATAGAGTCTCGGCCCAAAACCCCTAC	1323
Qy	6132	TAACTGTTGGATATGCCCTCCCTCGAACTTCARGCCCATATGTTTCAATCCCTGACTGTA	6191
Db	1324	TAACTGTTGGATATGCCCTCCCTCGAACTTCARGCCCATATGTTTCAATCCCTGACTGTA	1383
Qy	6192	ACAATGAAACAACCTTCAGCAAGAAAAATAACACCACTTCCGTTTATTAGTAGGACCTCTTGT	6251
Db	1384	ACAATGAAACAACCTTCAGCAAGAAAAATAACACCACTTCCGTTTATTAGTAGGACCTCTTGT	1443
Qy	6252	TTCAATTSAGAAATAACCCATACCTCAAAACCTCACTCGTTGTAATAATTTAGCAATACTAC	6311
Db	1444	TTCAATTSAGAAATAACCCATACCTCAAAACCTCACTCGTTGTAATAATTTAGCAATACTAC	1503
Qy	6312	ATACAAACAACCTCCCAATGCATCAGGTGGGTAACTCCTCCACACAAATAGTCTGCCT	6371
Db	1504	ATACAAACAACCTCCCAATGCATCAGGTGGGTAACTCCTCCACACAAATAGTCTGCCT	1563
Qy	6372	ACCTCAGGAATATTTTTTGTCTGTGTGACTCTCAGGCCTATCGTTGTTGTAATGGCTCTTC	6431
Db	1564	ACCTCAGGAATATTTTTTGTCTGTGTGACTCTCAGGCCTATCGTTGTTGTAATGGCTCTTC	1623
Qy	6432	AGAACTATGTGCTCTCTCATCTTTAGTGCCCCCYATGRCATCTACACTGAACAAGA	6491
Db	1624	AGAACTATGTGCTCTCTCATCTTTAGTGCCCCCYATGRCATCTACACTGAACAAGA	1683
Qy	6492	TTTTATACAGTTATGTCATATCTAAGCCCGCAACAAAGAGTACCACTTCTTCTTTTGT	6551
Db	1684	TTTTATACAGTTATGTCATATCTAAGCCCGCAACAAAGAGTACCACTTCTTCTTTTGT	1743
Qy	6552	TATAGGACGAGTGTGTAGTGCACTAGGTACTGGCATTTGGCGGTATCACACCTCTAC	6611
Db	1744	TATAGGACGAGTGTGTAGTGCACTAGGTACTGGCATTTGGCGGTATCACACCTCTAC	1803
Qy	6612	TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGAATGGAACGGGTCCGACCTC	6671
Db	1804	TCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGAATGGAACGGGTCCGACCTC	1863
Qy	6672	CCTGGTCACTTCAGATCACTTAACTCCCTAGCAGCAGTAGTCTCTTCAAATCGAAG	6731
Db	1864	CCTGGTCACTTCAGATCACTTAACTCCCTAGCAGCAGTAGTCTCTTCAAATCGAAG	1923
Qy	6732	AGCTTTAGACTCTGCTAAACCGCTGARGAGGGGAACTGTTTATTTTATAGGGGAAGATG	6791
Db	1924	AGCTTTAGACTCTGCTAAACCGCTGARGAGGGGAACTGTTTATTTTATAGGGGAAGATG	1983
Qy	6792	CTGTTATTATGTTAAATCAATCCGGAAATCGTCACTGAGAAAGTTTAAAGAAATTCAGATCG	6851



```

|||||
1984 CTGTTATTATGTTAATCAATCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCG 2043
QY
6852 AATACAAAGTAKAGCAGAGAGAGCTTCGAAACACTGAGACCTGGGCGCTCCTCAGCCRAATG 6911
Db
2044 AATACAAAGTAKAGCAGAGAGAGCTTCGAAACACTGAGACCTGGGCGCTCCTCAGCCRAATG 2103
QY
6912 GATGCCCTGGATCTCCCTCTTTAGGACCTCTAGCAGCTATAATATTGCTACTCTCTCTT 6971
Db
2104 GATGCCCTGGATCTCCCTCTTTAGGACCTCTAGCAGCTATAATATTGCTACTCTCTCTT 2163
QY
6972 TGGACCTGTATCTTTTACCTCTCTTTTAACTTTCTCTTCCAGAAATCGAAGCTGTAA 7031
Db
2164 TGGACCTGTATCTTTTACCTCTCTTTTAACTTTCTCTTCCAGAAATCGAAGCTGTAA 2223
QY
7032 ACTACAAATGGAGCCCAAGATGCACTCCAAAGACTAAGATCTACCGCAGACCCCTGGACCG 7091
Db
2224 ACTACAAATGGAGCCCAAGATGCACTCCAAAGACTAAGATCTACCGCAGACCCCTGGACCG 2283
QY
7092 GCCTGTAGCCCAAGATCTGATTTAAATGACATCAAAAGCAGCCCTCTGAGGAAATCTC 7151
Db
2284 GCCTGTAGCCCAAGATCTGATTTAAATGACATCAAAAGCAGCCCTCTGAGGAAATCTC 2343
QY
7152 ACCTGCACAACTCTACTACGCCCAATTTACGACGAAAGCACTTAGAGCGTSGTCGCC 7211
Db
2344 ACCTGCACAACTCTACTACGCCCAATTTACGACGAAAGCACTTAGAGCGTSGTCGCC 2403
QY
7212 AACCTCCCCCAACAGCACTTAGGTTTTCTGTTGAGATGGGGACTGAGAGACAGGACTAG 7271
Db
2404 AACCTCCCCCAACAGCACTTAGGTTTTCTGTTGAGATGGGGACTGAGAGACAGGACTAG 2463
QY
7272 CTGGATTTCTTAGGCTGATTAAGATTCCTTAAGACCTTAGTGAGAGGAGGACCAATCCAC 7331
Db
2464 CTGGATTTCTTAGGCTGATTAAGATTCCTTAAGACCTTAGTGAGAGGAGGACCAATCCAC 2523
QY
7332 CTTTAAACAGCGGGCTTGAATCTAGTACACCTGACCAATCAGAGAGCTCACTAAAT 7391
Db
2524 CTTTAAACAGCGGGCTTGAATCTAGTACACCTGACCAATCAGAGAGCTCACTAAAT 2583
QY
7392 GCTAATTAGGCAAGACAGAGGTTAAAGAAATAGCAATCATATTTGTTGAGAGGACCA 7451
Db
2584 GCTAATTAGGCAAGACAGAGGTTAAAGAAATAGCAATCATATTTGTTGAGAGGACCA 2643
QY
7452 GCAGAGGACCAATGATCGGGATATAAACCCAAAGTTCGAGCGGCAACGGCAACCCCC 7511
Db
2644 GCAGAGGACCAATGATCGGGATATAAACCCAAAGTTCGAGCGGCAACGGCAACCCCC 2703
QY
7512 TTTGGGTCCCTC 7524
Db
2704 TTTGGGTCCCTC 2716

```

```

RESULT 6
LOCUS BX365066/c
DEFINITION BX365066 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1046Y18 3-PRIME, mRNA sequence.
ACCESSION BX365066
VERSION BX365066.2 GI:46304105
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1071)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 5, 2003 this sequence version replaced gi:30374869.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

```

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?c=CS1A1012ZC10NPI&c=4215.r.

#### FEATURES

source

1. .1071

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clones="CS0D1046Y18"

/tissue\_type="PLACENTA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

#### ORIGIN

```

Query Match      12.9%; Score 969; DB 4; Length 1071;
Best Local Similarity 95.8%; Pred. No. 1.2e-272;
Matches 985; Conservative 25; Mismatches 16; Indels 2; Gaps 2;

QY 6520 CGCAACAAAGAGTACCCATTTCTTCTTTTGTATAGGAGCAGGAGTCTAGTGACACTA 6579
Db 1026 CCCSCACAAAGAGWCCCATVCTCCYTTTATATAGGAGCAGGAGTGCTA-GTGCACTA 968
QY 6580 GGTACTGCATTTGGCGGTATCACACCTCTACTCAGTTCTTACTACAAACTATCTCAAGAA 6639
Db 967 GGGACTGGCAATGGCGGTATAACACCTCTACTCAGTTCTTACTACAAACTATCTCAAGAA 908
QY 6640 CTAAATGGGACATGGAACGGGTGCCGACTCCCTGGTCACTTCCAAAGATCAACTTAAC 6699
Db 907 CTAAATGGGACATGGAACGGGTGCCGACTCCCTGGTCACTTCCAAAGATCAACTTAAC 848
QY 6700 TCCCTAGCAGCAGTAGTCTTTTAAATCGAAGAGCTTTAGACTYCTTAAACCCGATCAGA 6759
Db 847 TCCCTAGCAGCAGTAGTCTTTTAAATCGAAGAGCTTTAGACTYCTTAAACCCGATCAGA 788
QY 6760 GGGGGAACCTGTTTATTTTATAGGGAAGATGCTGTTTATTTATTTATTTATTTATTTAT 6819
Db 787 GGGGGAACCTGTTTATTTTATAGGGAAGATGCTGTTTATTTATTTATTTATTTATTTAT 728
QY 6820 GTCACTGAGAAAGTTTAAAGAAATTCAGATCGAATACAACTGAGAGGAGGAGCTTCA 6879
Db 727 GTCACTGAGAAAGTTTAAAGAAATTCAGATCGAATACAACTGAGAGGAGGAGCTTCA 668
QY 6880 AACACTGACCTGGGGCTCTCTCAGCCTATGAGATGAGGCTGAGGCTTCCCTTCTTAGGA 6939
Db 667 AACACTGACCTGGGGCTCTCTCAGCCTATGAGATGAGGCTGAGGCTTCCCTTCTTAGGA 608
QY 6940 CCTCTAGCAGCTATAATATTGCTACTCTCTTTTGGACCTGCTATCTTTTACCTCTCTGTT 6999
Db 607 CCTCTAGCAGCTATAATATTGCTACTCTCTTTTGGACCTGCTATCTTTTAACTCTCTGTT 548
QY 7000 AACTTTTGTCTTTCCAGAAATCGAAGCTGTAAATCTACAAATCGAGCCCAAGATGCACTCC 7059
Db 547 AACTTTTGTCTTTCCAGAAATCGAAGCTGTAAATCTACAAATCGAGCCCAAGATGCACTCC 488
QY 7060 AAGACTAAGATCTACCGCAGACCCCTGAGACCGGCTGTGTAGCCGACGATCTGATGTTAAT 7119
Db 487 AAGACTAAGATCTACCGCAGACCCCTGAGACCGGCTGTGTAGCCGACGATCTGATGTTAAT 428
QY 7120 GACATCAAGGACACCCCTCTCTGAGGAAATCTCAGCTGCAACACTCTACTAGCCCAAT 7179
Db 427 GACATCAAGGACACCCCTCTCTGAGGAAATCTCAGCTGCAACACTCTACTAGCCCAAT 368
QY 7180 TCACAGGAAGCAGTTAGAGCGGTGTGCGCCAACTCCGCCCAACAGCACTTAGTTTTC 7239

```



Db 367 TCAGCAGGAAGCAGTTAGAGGGTCTGTCGGCCAACTCCCAACAGCAGCTTAGGTTTCC 308

QY 7240 TGTTCAGATCGGGGACTGAGAGACAGACTAGCTAGCTGAGTTTCTTAGGCTGATTAAGATCC 7299

Db 307 TGTTCAGATCGGGGACTGAGAGACAGACTAGCTAGCTGAGTTTCTTAGGCTGATTAAGATCC 248

QY 7300 YTAAGCCTAGTGTGGAGGTGACACATCCACTTTAAACACGGGGCTTCAACTTAGYT 7359

Db 247 CTAAAGCTAGCTGGAGGTGACACATCCACTTTAAACACGGGGCTTCAACTTAGYT 188

QY 7360 CACACCTGACCAATCAGAGAGCTCACTAAATGCTAATTAGGCAAGACAGAGGTAAG 7419

Db 187 CACACCTGACCAATCAGAGAGCTCACTAAATGCTAATTAGGCAAGACAGAGGTAAG 128

QY 7420 AATAGCCATCATYTTATTCMTGAGAGCAGACAGAGGAGCAATGATCGGATATAA 7479

Db 127 AATAGCCATCATCTATTGCTTGAAGCAGACAGAGGAGCAATGATCGGATATAA 68

QY 7480 CCCAAGTTCGAGCGGCAACGGCAACCCCTTTGGGTCCTCCCTTTGATGGAGC 7539

Db 67 CCCAAGTTCGAGCGGCAACGGCAACCCCTTTGGGTCCTCCCTTTGATGGAGC 9

QY 7540 TCTGTTTT 7547

Db 8 TCKGTTTT 1

## RESULT 7

BX337769/c

LOCUS

DEFINITION BX337769 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0D1051Y13 3-PRIME, mRNA sequence.

ACCESSION

BX337769

VERSION

BX337769.2

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 998)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On May 2, 2003 this sequence version replaced gi:30337641.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0D1051AG07NP1&amp;c=4215.r.

Location/Qualifiers

1. 998

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0D1051Y13"

/tissue type="PLACENTA COT 25-NORMALIZED"

/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

division of Invitrogen. This sequence belongs to sequence cluster

4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0D1051AG07NP1&amp;c=4215.r.

Location/Qualifiers

1. 998

/organism="Homo sapiens"

Query Match

Best Local Similarity

12.3%; Score 921.2; DB 4; Length 998;

95.8%; Pred. No. 1.5e-258;

ORIGIN

QY

Db

QY

Matches 956; Conservative 26; Mismatches 11; Indels 5; Gaps 4;

QY 6559 GCAGGAGTGTCTAGTGCCTAGTACTGGCATTTGGCGGTATCAAAACC-TCTACTCAGTT 6617

Db 998 GCAGGAGTGTCTAGTGCCTAGTACTGGCATTTGGCGGTATCAAAACC-TCTACTCAGTT 939

QY 6618 CTACTCAAACTATCTCAAGAACT-AAATGGGACATAGAAACGGGTGCGGACTCCCTGG 6676

Db 938 CTACTCAAACTATCTCAAGAACT-AAATGGGACATAGAAACGGGTGCGGACTCCCTGG 879

QY 6677 TCACCTTGCAGAGATCAACTTAACTCCCTAGCAGCAGTGTCTTCAAATTCGAAGAGCTT 6736

Db 878 TCACCTTGCAGAGATCAACTTAACTCCCTAGCAGCAGTGTCTTCAAATTCGAAGAGCTT 819

QY 6737 TAGACTYGTAAACCGCTGAPAGAGGGGAACTGTGTTATTTTATAGGGGAGAAATGCTGTT 6796

Db 818 TAGACTYGTAAACCGCTGAPAGAGGGGAACTGTGTTATTTTATAGGGGAGAAATGCTGTT 759

QY 6797 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAATTCGAGATCGAATAC 6856

Db 758 ATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAATTCGAGATCGAATAC 699

QY 6857 AAGCTAGACAGAGAGCTTCGAAACACCTGGAACCTGGGGCCCTCCTCAGCCATGATGTC 6916

Db 698 AAGCTAGACAGAGAGCTTCGAAACACCTGGAACCTGGGGCCCTCCTCAGCCATGATGTC 639

QY 6917 CCTGGATTCCTCCCTTCTTAGGACCTCTAGCAGCTATAATATTTGCTACTCTCTTTGGAC 6976

Db 638 CCTGGATTCCTCCCTTCTTAGGACCTCTAGCAGCTATAATATTTGCTACTCTCTTTGGAC 579

QY 6977 CCTGTATCTTTTACCTCTCTTTTAACTTTGCTCTTCCAGAAATCGAAGTGTAAACTAC 7036

Db 578 CCTGTATCTTTTAACTCTCTTTTAACTTTGCTCTTCCAGAAATCGAAGTGTAAACTAC 519

QY 7037 AATGGAGCCCAAGATGAGTCCAGACTCAAGTCTACCGCAGACCCCTGGACCGGCTG 7096

Db 518 AATGGAGCCCAAGATGAGTCCAGACTCAAGTCTACCGCAGACCCCTGGACCGGCTG 459

QY 7097 YTAGCCCAAGCTGATGTTAATGACATCAAA-GGCACCCCTCCTCAGAGAAATCTCAGCT 7155

Db 458 YTAGCCCAAGCTGATGTTAATGACATCAAA-GGCACCCCTCCTCAGAGAAATCTCAGCT 399

QY 7156 GCACAACTCTACTACGCCCCAAATTCAGCAGGAGCAGTTAGAGCGGTGTCGCGCAACC 7215

Db 398 GCACAACTCTACTACGCCCCAAATTCAGCAGGAGCAGTTAGAGCGGTGTCGCGCAACC 339

QY 7216 TCCCAACAGCAGCTTAGGTTTCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGG 7275

Db 338 TCCCAACAGCAGCTTAGGTTTCTGTTGAGATGGGGGACTGAGAGACAGGACTAGCTGG 279

QY 7276 ATTTCTAGGCTGATTAAGAAATCCYTAAGCCTAGTGGGAGGTGACCACTCCACTTTT 7335

Db 278 ATTTCTAGGCTGATTAAGAAATCCYTAAGCCTAGTGGGAGGTGACCACTCCACTTTT 219

QY 7336 AAACAAGGCTTGCAGCTTAGCTTCACTCACTCACTCACTCACTCACTCACTCACTCACT 7395

Db 218 AAACAAGGCTTGCAGCTTAGCTTCACTCACTCACTCACTCACTCACTCACTCACTCACT 159

QY 7396 ATTAGCAAGACAGAGGAGTAAAGAAATAGCCAAATCATTTATTTGCTGAGAGCAGAG 7455

Db 158 ATTAGCAAGACAGAGGAGTAAAGAAATAGCCAAATCATTTATTTGCTGAGAGCAGAG 99

QY 7456 GAGGAGCAATGATCGGAGTATAAACCAGTGTTCGAGCCGCGCAAGGCAACCCCTTTG 7515

Db 98 GAGGAGCAATGATCGGAGTATAAACCAGTGTTCGAGCCGCGCAAGGCAACCCCTTTG 39

QY 7516 GGTCCC--CTCCCTTTGATGGGAGCTCTGTTTTCATG 7551

Db 38 GGTCCCNCCCTTTGTANTGGGAGTCTGTTTTCATG 1

RESULT 8

BX378303/c



DB  
121 GATGGGGGTGGAGTTCAGGATCAGGCAGAGAAAAAATCATGTAAGAAGTAATCTCCCAA 180



QY 7240 TGTGAGTGGGAGTGGAGAGCAGGACTAGCTGGATTTCCTAGGCTGATTAAGATCC 7299  
 Db 315 TGTGAGTGGGAGTGGAGAGCAGGACTAGCTGGATTTCCTAGGCTGATTAAGATCC 256  
 QY 7300 YTAAGCTAGTGGGAGTGGAGAGCAGGACTAGCTGGATTTCCTAGGCTGATTAAG 7359  
 Db 255 CTAAGCTAGTGGGAGTGGAGAGCAGGACTAGCTGGATTTCCTAGGCTGATTAAG 196  
 QY 7360 CACACCTGACCAATCAGAGAGCTCACTAAATGCTAAATGCTAAATGCTAAATGCTAA 7419  
 Db 195 CACACCTGACCAATCAGAGAGCTCACTAAATGCTAAATGCTAAATGCTAAATGCTAA 136  
 QY 7420 AATAGTCCCAATCATYATTTCCTGAGAGCAGCAGGAGGACCAATGATCGGATATAA 7479  
 Db 135 TAATKNVRATTATTATTCCTGAGAGCAGCAGGAGGACCAATGATCGGATATAA 76  
 QY 7480 CCAAGTTCGAGCGGCAACCGCAACCGCTTTGGTCCCTCCCTTTGTTATGGAGC 7539  
 Db 75 YCCAAGTTCGAGCGGCAACCGCAACCGCTTTGGTCCCTCCCTTTGTTATGGAGC 17  
 QY 7540 TCTGTTTTCATGCTAT 7555  
 Db 16 TTTTNTTNTTATTT 1

## RESULT 12

LOCUS BX378304 1014 bp mRNA linear EST 26-APR-2004  
 DEFINITION BX378304 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 clone CS0D1018YH07 5-PRIME, mRNA sequence.

ACCESSION BX378304  
 VERSION BX378304.2 GI:46569716

KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.

REFERENCE 1 (bases 1 to 1014)  
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On May 8, 2003 this sequence version replaced gi:30443114.

Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 4215.r

For more information about this cluster, see  
<http://www.genoscope.cns.fr/cdna?8=CS0D1018YH07&4215.r>.

## FEATURES

source  
 1. .1014  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1018YH07"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was  
 sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 11.6%; Score 866.6; DB 4; Length 1014;  
 Best Local Similarity 95.0%; Pred. No. 1.8e-242;  
 Matches 874; Conservative 24; Mismatches 21; Indels 1; Gaps 1;

QY 80 TGTTCATGCTATTATTCATCTATTATAATCTTGCAACGTGCACTCTCTGCTGCAATGTTTC 139  
 Db 1 TGTTCATGCTATTATTCATCTATTATAATCTTGCAACGTGCACTCTCTGCTGCAATGTTTC 60  
 QY 140 TTACGGCTCGAGCTGAGCTTTTGTCTCACCGTCCACCACTGCTCTTTGGCCACCAACCGCANA 199  
 Db 61 TTACGGCTCGAGCTGAGCTTTTGTCTCACCGTCCACCACTGCTCTTTGGCCACCAACCGCAGA 120  
 QY 200 CTTGCCGCTGACTCCCATCCCTCTGGATCTGTCAGGGTGTCCGCTGTGCTCTCTGATCCAG 259  
 Db 121 CTTGCCGCTGACTCCCATCCCTCTGGATCTGTCAGGGTGTCCGCTGTGCTCTCTGATCCAG 180  
 QY 260 CGARGCGCCCATTCGCGCTCCCAATTTGGGCTAAAGCTTGGCCATTGCTGCTGCTGACCGCTA 319  
 Db 181 CGARGCGCCCATTCGCGCTCCCAATTTGGGCTAAAGCTTGGCCATTGCTGCTGCTGACCGCTA 240  
 QY 320 AGTGCCTGGGTTTGTCTTAATTTGAGCTGAACACTANTCACTGGGTTCCATGGTTCCTCTTC 379  
 Db 241 AGTGCCTGGGTTTGTCTTAATTTGAGCTGAACACTANTCACTGGGTTCCATGGTTCCTCTTC 300  
 QY 380 TTGACCCACCGGCTTCTAATAKAACTATAACACTTACCACATGGGCCCAAGATTCCATTTCC 439  
 Db 301 TTGACCCACCGGCTTCTAATAKAACTATAACACTTACCACATGGGCCCAAGATTCCATTTCC 360  
 QY 440 TTGGAATCCGTGAGGSCAAGAACTCCAGGTCAGAGATACAGAGCTTGGCCACCATCTTG 499  
 Db 361 TTGGAATCCGTGAGGSCAAGAACTCCAGGTCAGAGATACAGAGCTTGGCCACCATCTTG 419  
 QY 500 GAAGCGGCTGCTACGCTCTCGAGTGGTTCACCACTCTTGGGAGCTCTGTGAGCAA 559  
 Db 420 GAAGCGGCTGCTACGCTCTCGAGTGGTTCACCACTCTTGGGAGCTCTGTGAGCAA 479  
 QY 560 GGACCCCGGCTTACATTTTGGCRACCAACGAGCATCCMAAGTGATGGGAAACGTTTC 619  
 Db 480 GGACCCCGGCTTACATTTTGGCRACCAACGAGCATCCMAAGTGATGGGAAACGTTTC 539  
 QY 620 CCGCGAAGCAAAAACGCGCTTAAAGACGTATTTCTGGARAAATTTGGGAMCAATTTGACCCCTC 679  
 Db 540 CCGCGAAGCAAAAACGCGCTTAAAGACGTATTTCTGGARAAATTTGGGAMCAATTTGACCCCTC 599  
 QY 680 AGACACTAAGAAAGAAACGACTTATATTTCTGCAAGTGGCGCTGGCACTCTGTAGGGA 739  
 Db 600 AGACACTAAGAAAGAAACGACTTATATTTCTGCAAGTGGCGCTGGCACTCTGTAGGGA 659  
 QY 740 AGTATAAATTATAACACCATCTTACAGCTAGACCTCTTTGTAGAAAAGGCAAAATCGAGT 799  
 Db 660 AGTATAAATTATAAAGGCAAAATTTGCAAGGCAAAATTTGCAAGGCAAAATCGAGT 719  
 QY 800 GAAGTCCCATAGTACAAACTTTCTTTTCAATTAAGAGACAACTCACAATTTATGTAAAAAG 859  
 Db 720 GAAGTCCCATAGTACAAACTTTCTTTTCAATTAAGAGACAACTCACAATTTATGTAAAAAG 779  
 QY 860 TGTGATTATGCTCTACAGGAAGCTTACAGTCTACCTTCCCTATCCCAAGCATTCCTCCGAC 919  
 Db 780 TGTGATTATGCTCTACAGGAAGCTTACAGTCTACCTTCCCTATCCCAAGCATTCCTCCGAC 839  
 QY 920 TCTTCTCCCACTAATAGGACCCCTTCAACCCCAATTTGTCCTCAAAAGGAGATAGACAA 979  
 Db 840 TCTTCTCCCACTAATAGGACCCCTTCAACCCCAATTTGTCCTCAAAAGGAGATAGACAA 899  
 QY 980 AAGGGTAAACAGTGAACCAA 999  
 Db 900 AAGGGTAAACAGTGAACCAA 919

RESULT 13  
 LOCUS CR749521

DEFINITION Homo sapiens mRNA; cDNA DKFZp686l21167 (from clone DKFZp686l21167).  
 ACCESSION CR749521  
 VERSION CR749521.1 GI:51476727  
 KEYWORDS HTC.  
 SOURCE Homo sapiens (human)

CR749521 6400 bp mRNA linear HTC 19-AUG-2004  
 Homo sapiens mRNA; cDNA DKFZp686l21167 (from clone DKFZp686l21167).  
 CR749521  
 CR749521.1 GI:51476727  
 HTC.  
 Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.

REFERENCE 1 (bases 1 to 6400)  
 Koehrer, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,  
 Fobo, G., Han, M. and Wiemann, S.  
 The German cDNA Consortium

CONSRTH  
 TITLE Direct Submission  
 JOURNAL Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764  
 Neuherberg, GERMANY

COMMENT  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 sequenced by BMFZ (Biomedical Research Center at the  
 Heinrich-Heine-University, Dueseldorff/Germany) within the cDNA  
 sequencing consortium of the German Genome Project. This clone  
 (DKFZp686121167) is available at the RZPD Deutsches  
 Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.  
 Please contact RZPD for ordering:  
 http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp686121167  
 Further information about the clone and the sequencing project is  
 available at http://mips.gsf.de/projects/cdna/.

FEATURES  
 Location/Qualifiers  
 1..6400  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="DKFZp686121167"  
 /tissue\_type="uterus endometrium, primary cell culture"  
 /clone\_lib="686 (synonym: hicc3). Vector pSport1\_Sfi; host  
 DH10B; sites SfiIA + SfiIB"  
 /dev\_stage="adult"  
 /note="hypothetical protein"  
 1..6400  
 /gene="DKFZp686121167"  
 209..1573  
 /gene="DKFZp686121167"  
 /codon\_start=1  
 /product="hypothetical protein"  
 /protein\_id="CAH18335.1"  
 /db\_xref="GI:51476728"  
 /translation="MTWAAASTCAALLILWLTGDLKVMAGGTQITPLNDNV  
 TIFNIFSQPIQLTSMGITFWMKSLTFDKEVKVFEFFGDHQAEPFCAIVSPWRLKS  
 GDASIRLQILEAGEVCEVAVTPPLAQGTVOLEVVASPASRLLLDQVGMKEDK  
 YNCESGYPEALNITWETQKPHPIEISDVITGTIKMDGTFFNTSCLKLNS  
 QBDPTGVQVVRHSLHTPLRSNFTLFAARHSLSETKDNFSLHWPPIFGVGLV  
 LLIVLPWKIKCNKSSSAYTPLKILKHNSFDOTLKEHLIFFTCRAMPFYSIQDQ  
 EAWPEGSVINIIOQLDVCROGKWEVYVQAFALRNPDLCQCCRIDPALLTV  
 TSGKSIDDNSTKSEQTPREHSDAVDPILPVSPWPPPTTTTFLVSSQRPPTLL  
 LPLQ"

gene  
 1..6400  
 /gene="DKFZp686121167"  
 209..1573  
 /gene="DKFZp686121167"  
 /codon\_start=1  
 /product="hypothetical protein"  
 /protein\_id="CAH18335.1"  
 /db\_xref="GI:51476728"  
 /translation="MTWAAASTCAALLILWLTGDLKVMAGGTQITPLNDNV  
 TIFNIFSQPIQLTSMGITFWMKSLTFDKEVKVFEFFGDHQAEPFCAIVSPWRLKS  
 GDASIRLQILEAGEVCEVAVTPPLAQGTVOLEVVASPASRLLLDQVGMKEDK  
 YNCESGYPEALNITWETQKPHPIEISDVITGTIKMDGTFFNTSCLKLNS  
 QBDPTGVQVVRHSLHTPLRSNFTLFAARHSLSETKDNFSLHWPPIFGVGLV  
 LLIVLPWKIKCNKSSSAYTPLKILKHNSFDOTLKEHLIFFTCRAMPFYSIQDQ  
 EAWPEGSVINIIOQLDVCROGKWEVYVQAFALRNPDLCQCCRIDPALLTV  
 TSGKSIDDNSTKSEQTPREHSDAVDPILPVSPWPPPTTTTFLVSSQRPPTLL  
 LPLQ"

CDS  
 209..1573  
 /gene="DKFZp686121167"  
 /codon\_start=1  
 /product="hypothetical protein"  
 /protein\_id="CAH18335.1"  
 /db\_xref="GI:51476728"  
 /translation="MTWAAASTCAALLILWLTGDLKVMAGGTQITPLNDNV  
 TIFNIFSQPIQLTSMGITFWMKSLTFDKEVKVFEFFGDHQAEPFCAIVSPWRLKS  
 GDASIRLQILEAGEVCEVAVTPPLAQGTVOLEVVASPASRLLLDQVGMKEDK  
 YNCESGYPEALNITWETQKPHPIEISDVITGTIKMDGTFFNTSCLKLNS  
 QBDPTGVQVVRHSLHTPLRSNFTLFAARHSLSETKDNFSLHWPPIFGVGLV  
 LLIVLPWKIKCNKSSSAYTPLKILKHNSFDOTLKEHLIFFTCRAMPFYSIQDQ  
 EAWPEGSVINIIOQLDVCROGKWEVYVQAFALRNPDLCQCCRIDPALLTV  
 TSGKSIDDNSTKSEQTPREHSDAVDPILPVSPWPPPTTTTFLVSSQRPPTLL  
 LPLQ"

ORIGIN  
 Query Match 11.3%; Score 848.8; DB 6; Length 6400;  
 Best Local Similarity 54.9%; Pred. No. 8.3e-237;  
 Matches 2156; Conservative 112; Mismatches 1531; Indels 130; Gaps 33;  
 1057 GGCACGACAGAGTGATGCTGCTTTTCTCCAGACTTAAGCAATAAAGACAC 1116  
 1589 GGTGCCACTAGGCTCAAGTCCCTTTTCAATACAGGACCTTGGGCAATAAGAGGGAC 1648  
 1117 TTAGGTAAATCTCAGATAAATCTGATGCTATATTGTTTCAAGGTTTAGACAA 1176  
 1649 CTGGGCAAGTCTCTGATGA-CTGATAGATATAGAGGCTTTCAAAACCTAATCTAG 1707  
 1177 TTCTTTGATCTGACATGGAGAGATATATGTCATGCTTAATACAGACATAACCCCAA 1236  
 1708 GTGTTTGACTTACTTGGGGTGATGT-TATGTTGCTCTTAAACTCTTAACACTACAGAG 1766  
 1237 TGAGAGAAGTGCCACCA--TAATGCGAGCTGAGGTGTTGGGACTCTGGTATCTCAGT 1294  
 1767 AAACAGGTAGCCCTGAGGACAGCAAAATTCAGAGATGAACAACATGCTCTCTATAGC 1826  
 1295 CAGGTCAATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1354

Db 1827 CACTCAATAAGGAAGAAGGTGAGAAAACAACCCAGCTTTCATATACCCAAATAGGAAGCGAG 1886  
 QY 1355 GCAGTTCAGTCTTASACCTCATTTGGGGACACAGAAATCAATGTAACATGGAGATTTGGTG 1414  
 Db 1887 GCAGTGGCCCTTGAACAACATAATTCGGACCCAG----TGGTACCTTTAGATGAGTGCA 1942  
 QY 1415 CTGCAGACATTTGCTAACTTGTGTCTASAGGACTAAGGAAAACCTASGAGAAGAAATCTTA 1474  
 Db 1943 AAGAAACATTTTCAATATGCTTATTAGAGGCTTACAGAGGACCA-GAGCCAGCCCTC 2001  
 QY 1475 YGAATTACTCAATGATGTCCACCAATAACACAGGGGAGGAGGAGAAATCTTACTGCTTTT 1534  
 Db 2002 TTAATTACCCCAAACTCTCCATGATAGATCA-AAAACTAGATGAAATCCCTCAGCCCTT 2060  
 QY 1535 CTGAGAGACATAAGGGAGGCATTTAGGAAGCGTCTCTCTGTCACTGT-----1583  
 Db 2061 TTGGAAGGCTGAGAGAGGCTGTAGTGAACAACACTTCTCTGTGTGAGGTATGTT 2120  
 QY 1584 -----ACTCTTCTGAGGCGCAACTAATCTTAAAGCGTAAAGTTTATCAC 1626  
 Db 2121 TATACCTCACATAAATTAATTCAGTTGAGGACAGCTAATTTTGAAGCAAGTTTATTATC 2180  
 QY 1627 TCAGTTCAGCTGCAGACATTTAGAA-AAAACTTCAAAAGTCTCCGCTAGGCCCGGAGCAAAA 1685  
 Db 2181 TCAGACAGCCCTGATATCAGAGGAAAGTCTAGAGAAAACAGGCCATGGGATTTAGATAGCAC 2240  
 QY 1686 CTTAGAAACCTTATTAACCTTGGCAACCTCGGTTTTTATTAATAGATAGAGGAGGACGA 1745  
 Db 2241 CTTACAAACCTCTCCTGAGGTAGCCACCTCGTCTTTTATTAATAGGACAGGAGGAGGC 2300  
 QY 1746 GGCAGAACAGGACAAAACGGGATTAATAAAGGACCCGCTTTAGTCATGACCCCTCAGGC 1805  
 Db 2301 CCAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2333  
 QY 1806 AAGTGGACTTTGGAGGCTCTGAGAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1865  
 Db 2334 -AGAGACTCTGTTGGCCACTTTTGTATATAAACCAGTATCCCCAAGGTACACCTGCTAG-- 2390  
 QY 1866 CTTGCTTCAGTGGCTCTACAGGACACTTTTAAAGAGATGTCCTCAAGTAGAGTAAGTAAGC 1925  
 Db 2391 -TTGCTACCAATGTGGCAAACTAGGGCACTTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGG 2449  
 QY 1926 GGCCTCTGTCCTGCTGCTTATTTCAAGGAGTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGG 1985  
 Db 2450 AGCACCTTCAACCTGTCAGTATACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2509  
 QY 1986 GACAAAGGTCTTTGAGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2042  
 Db 2510 AGCATAGTCTGAGTCCAGAGCCAGTCTCACAATGGTCCAGGAGGAGGAGGAGGAGGAGG 2569  
 QY 2043 -----GGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2093  
 Db 2570 TGGGGCTCAACTCCCTGGCTCCAGTGGCTCAGACTGCTGTTACTACCCGGGAGGAGGAGGAGG 2629  
 QY 2094 TATGCTTGACCATTTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2150  
 Db 2630 TGATTCTGAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2688  
 QY 2151 TAGTCTTACTCTTCTGCTCCGAGCAAACTGCTCTCCAGATCTGCTCACTATTCTGAGGGGGT 2210  
 Db 2689 CTGTTTCTCTCCAAATTCAGGCCCTTCGTTCTCTCC--CATAGCAACAACCATGATGGGCA 2745  
 QY 2211 CCAATAGAGGGGAGTCACTAGATGATCTTTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2270  
 Db 2746 TCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2805  
 QY 2271 TTTATTTCTTTTCAATGCTTTTCTAAATGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2330  
 Db 2806 --TACTGTTTACAGGGCTTTTGTGATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2863  
 QY 2331 AGACATTTAGCAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2390







clone CS0D1051YM13 5-PRIME, mRNA sequence.

ACCESSION  
BX389656  
VERSION  
BX389656.1  
KEYWORDS  
GI:30462930  
SOURCE  
EST.  
ORGANISM  
Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

1 (bases 1 to 931)

Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

Full-length cDNA libraries and normalization  
Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0BAP0272B05\_AP02531\_1&c=4215.r

#### FEATURES

source

Location/Qualifiers

1..931

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0D1051YM13"

/tissue\_type="PLACENTA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

#### ORIGIN

Query Match 11.0%; Score 820.8; DB 4; Length 931;  
Best Local Similarity 97.1%; Pred. No. 5.5e-229;  
Matches 899; Conservative 8; Mismatches 12; Indels 7; Gaps 7;  
QY 5848 AAAATGATTAATCCTAGTTGTCTCTGGAGGACTTGGAGTCACTGCTGTGGACTTACTTTC 5907  
Db |||||  
2 AAAAGATTAATCCTAGTTGTCTCTGGAGGACTTGGAGTCACTGCTGTGGACTTACTTTC 61  
QY 5908 ACCCAAACTGGTATGTCTGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAACATGTA 5967  
Db |||||  
62 ACCCAAACTGGTATGTCTGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAACATGTA 121  
QY 5968 AAGAAGATTAATCCTCAAACTCACCGGGTACATGGCACCTCTAGCCCTACAAAGGACTA 6027  
Db |||||  
122 AAGAAGATTAATCCTCAAACTCACCGGGTACATGGCACCTCTAGCCCTACAAAGGACTA 181  
QY 6028 GATCTCTCAAACTACATGAACCCCTCCGTACCATCTCGCTGGTGAAGCCCTATTATAAT 6087  
Db |||||  
182 GATCTCTCAAACTACATGAACCCCTCCGTACCATCTCGCTGGTGAAGCCCTATTATAAT 241  
QY 6088 ACCACCTCTCACTGGGCTCCATGAGTCTCGGCCAAAACCCCTACTAACTGTTGGATATGC 6147  
Db |||||  
242 ACCACCTCTCACTGGGCTCCATGAGTCTCGGCCAAAACCCCTACTAACTGTTGGATATGC 301  
QY 6148 CTCCTCTGAACTTCGAGCCATATGTTCAATCCCTGTACCTGAACAATGGAACTTTC 6207  
Db |||||  
302 CTCCTCTGAACTTCGAGCCATATGTTCAATCCCTGTACCTGAACAATGGAACTTTC 361  
QY 6208 AGCACAAGATAAACCCTCTCCGTTTGTAGTGGACCTCTGTTTCCATSTGGAAATA 6267  
Db |||||  
362 AGCACAAGATAAACCCTCTCCGTTTGTAGTGGACCTCTGTTTCCATSTGGAAATA 421  
QY 6268 ACCCATACCTCAAACTCCTGTTGTAATAATTTAGCAATACTACATACCAACCACTCC 6327

Search completed: August 2, 2006, 01:42:15  
Job time : 30592 secs

Db 422 ACCCATACCTCAAACTCCTCCTGTGTAATAATTTAGCAATACTACATACCAACCACTCC 481  
QY 6328 CAATGCATCAGGTGGGTAACTCCTCCACACAATAATAGTCTGCTACCTCAGGAATATTT 6387  
Db 482 CAATGCATCAGGTGGGTAACTCCTCCACACAATAATAGTCTGCTACCTCAGGAATATTT 541  
QY 6388 TTTGCTGTGGTACCTCAAGCTATCGTTGTTGAATGGCTTTTCAAGATCTATGTGCTTC 6447  
Db 542 TTTGCTGTGGTACCTCAGCTATCGTTGTTGAATGGCTTTTCAAGATCTATGTGCTTC 601  
QY 6448 CTCTCATTTCTAGTGGCCCCYATGRCATCTACACTGAACAAGATTTATACAGTTATGTC 6507  
Db 602 CTCTCATTTCTAGTGGCCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTC 661  
QY 6508 ATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCTTTTGTATAGGAGCAGGAGTG 6567  
Db 662 ATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCTTTTGTATAGGAGCAGGAGTG 721  
QY 6568 CTAGGTGCATAGGTACTGGCATTTGGCGGTATCACAACTCTACTCAGTTCTACTACAAA 6627  
Db 722 CTAGGTGCATAGGTACTGGCATTTGGCGG-ATCACAACTCTACTCAGTTCTACTACANA 780  
QY 6628 CTATCTCAAGAACTAAATGGGAC-ATGGAAACGGGTGCGCGACT-CCCTGGTCACCTTGC 6685  
Db 781 CTATCTCAAGAACTAAATGGGACAAATGTAACGGGTGCGCGACTCCCTGGTCACCTTCG 840  
QY 6686 AAGATCAACTTAAC-TCCCTAGCAGCAGTA-GTCCCTTCRAAATCGAAG-AGCTTTAGACT 6742  
Db 841 CAGATCAACTTACCTTCCCTAGCAGCAGGAGGTCTTCAAAATCGAAGAGCTTTAGCCT 900  
QY 6743 YGCTAA-CGCTGAGAGAGGGGGAAC 6767  
Db 901 TGTAAACCCGCTGCAAGAGGGGGAAC 926

**This Page Blank (uspto)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2006, 17:12:25 ; Search time 1319 Seconds

(without alignments)  
10755.684 Million cell updates/sec

Title: US-10-717-580-11

Perfect score: 7489.6

Sequence: 1 caacaatcggtataataacc.....tattaatacttgactgtgr 7582

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/ina/1 COMB.seq:\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/ina/5 COMB.seq:\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6A COMB.seq:\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6B COMB.seq:\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/ina/7 COMB.seq:\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/ina/7 COMB.seq:\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/2/ina/7 COMB.seq:\*
- 8: /EMC\_Celerra\_SIDS3/ptodata/2/ina/7 COMB.seq:\*
- 9: /EMC\_Celerra\_SIDS3/ptodata/2/ina/7 COMB.seq:\*
- 10: /EMC\_Celerra\_SIDS3/ptodata/2/ina/7 COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6448	86.1	13537	3	US-09-949-016-17382 Sequence 17382, A
2	6436	85.9	10499	3	US-09-719-554-3 Sequence 3, Appli
3	6124.6	81.8	8523	3	US-09-573-080A-21 Sequence 21, Appl
4	6124.6	81.8	8523	5	US-09-854-867-21 Sequence 21, Appl
5	2994.2	40.0	7772	3	US-09-949-016-17417 Sequence 17417, A
6	2994.2	40.0	7797	3	US-09-949-016-12249 Sequence 12249, A
7	2888.6	38.6	2946	3	US-09-175-928-3 Sequence 3, Appli
8	2615.4	34.9	168394	3	US-09-949-016-13002 Sequence 13002, A
9	2550.4	34.1	2599	3	US-09-719-554-1 Sequence 1, Appli
10	2282.4	30.5	2784	3	US-09-719-554-4 Sequence 4, Appli
11	2220.4	29.6	145320	3	US-09-949-016-15858 Sequence 15858, A
12	2216	29.6	46340	3	US-09-719-554-66 Sequence 66, Appl
13	2204	29.4	2763	3	US-09-949-016-5640 Sequence 5640, Ap
14	2168	28.9	8399	3	US-09-573-080A-26 Sequence 26, Appl
15	2168	28.9	8399	5	US-09-854-867-26 Sequence 26, Appl
16	2032	27.1	2055	3	US-09-719-554-22 Sequence 22, Appl
17	1837.4	24.5	2304	3	US-08-979-847B-87 Sequence 87, Appl
18	1507.4	20.1	276687	3	US-09-949-016-13840 Sequence 13840, A
19	1464.6	19.6	1799	3	US-09-719-554-5 Sequence 5, Appli
20	1452.4	19.4	2364	3	US-08-979-847B-88 Sequence 88, Appl
21	1415.8	18.9	2391	3	US-08-691-563C-57 Sequence 57, Appl
22	1415.8	18.9	2391	3	US-08-374-766-57 Sequence 57, Appl
23	1415.8	18.9	2391	3	US-08-979-847B-53 Sequence 53, Appl

24	1341.8	17.9	8919	3	US-09-573-080A-54 Sequence 54, Appl
25	1341.8	17.9	8919	5	US-09-854-867-54 Sequence 54, Appl
26	1324.2	17.7	1481	3	US-08-979-847B-105 Sequence 105, App
27	1312.4	17.5	1600	3	US-08-979-847B-207 Sequence 207, App
28	1301.8	17.4	1600	3	US-08-979-847B-206 Sequence 206, App
29	1282.4	17.1	1597	3	US-08-979-847B-205 Sequence 205, App
30	1200.2	16.0	3910	3	US-09-120-653D-1 Sequence 1, Appli
31	1195.4	16.0	1326	3	US-09-719-554-2 Sequence 2, Appli
32	1130.6	15.1	1740	3	US-09-719-554-61 Sequence 61, Appl
33	1087.2	14.5	1511	3	US-08-979-847B-117 Sequence 117, App
34	1081.4	14.4	1774	3	US-09-719-554-17 Sequence 17, Appl
35	1069.4	14.3	1577	3	US-08-691-563C-89 Sequence 89, Appl
36	1069.4	14.3	1577	3	US-09-374-766-89 Sequence 89, Appl
37	1069.4	14.3	1577	3	US-08-979-847B-83 Sequence 83, Appl
38	996.8	13.3	2330	3	US-09-120-653D-4 Sequence 4, Appli
39	954.2	12.7	1080	3	US-09-719-554-28 Sequence 28, Appl
40	950	12.7	1489	3	US-09-719-554-6 Sequence 6, Appli
41	947.8	12.7	1167	3	US-08-691-563C-61 Sequence 61, Appl
42	947.8	12.7	1167	3	US-09-374-766-61 Sequence 61, Appl
43	947.8	12.7	1167	3	US-08-979-847B-57 Sequence 57, Appl
44	932.4	12.4	1158	2	US-08-471-724-1 Sequence 1, Appli
45	932.4	12.4	1158	2	US-08-471-969-1 Sequence 1, Appli

## ALIGNMENTS

### RESULT 1

US-09-949-016-17382  
; Sequence 17382, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17382  
; LENGTH: 13537  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-17382

Query Match	86.1%	Score 6448; DB 3; Length 13537;
Best Local Similarity	95.7%	Pred. No. 0;
Matches 6681; Conservative 195; Mismatches 50; Indels 53; Gaps 27;		
QY	606	GATGGGAACGTTCCCGCAACACACAAACGCGCCCTAACACGCTATTCTCGARAATTGGGA 665
Db	4610	GATGGGAACGTTCCCGCAACACACAAACGCGCCCTAACACGCTATTCTCGARAATTGGGA 4669
QY	666	MOAATTTGACCTCAGACACTAAGAAAGAACGACTTATATCTCTGCAGTCCGCGCTG 725
Db	4670	CCAAATTTGACCTCAGACACTAAGAAAGAACGACTTATATCTCTGCAGTCCGCGCTG 4729
QY	726	GCACCTCTGAGGAAGTATAAATTATACACCATCTTACAGCTAGACCTCTTTTGTAGAA 785
Db	4730	GCACCTCTGAGGAAGTATAAATTATACACCATCTTACAGCTAGACCTCTTTTGTAGAA 4789
QY	786	RAGGCAAAATGAGTGAAGTGCATTAAGTACAACTCTTTTTCATTAGAGCACTCAC 845
Db	4790	RAGGCAAAATGAGTGAAGTGCATTAAGTACAACTCTTTTTCATTAGAGCACTCAC 4849
QY	846	RAATATGAAAAAGTGTGATTATATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATC 905







Db 9180 CTTTAGCCAAATATCAACAAGTTCTTAAAAATTAACAAGAACCTTATCCCTGAGAAGAGG 9239  
Qy 5284 GAAAGAACTATTCCACCWGTGACATGGTATTAGTCAAGTCCCTTCYCTCTAAATGCC 5343  
Db 9240 GAAAGAACTATTCCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCTCTAAATGCC 9299  
Qy 5344 CATCCCTAGATACATCTCTGGGAAGACCTACCCAGTCAATTTATATACCCCACTGCGG 5403  
Db 9300 CATCCCTAGATACATCTCTGGGAAGACCTACCCAGTCAATTTATCTACCCCACTGCGG 9359  
Qy 5404 TTAAGTGGCTGGAGTCTTGGATACATCACATTTGAGTCAAAATCTCTGGATACCTG 5463  
Db 9360 TTAAGTGGCTGGAGTCTTGGATACATCACATTTGAGTCAAAATCTCTGGATACCTG 9419  
Qy 5464 CAAAGAACCTGAAATCCAGAGACAAAGCTAGTATTTCTGTGAACTCTTAGAGGATT 5523  
Db 9420 CAAAGAACCTGAAATCCAGAGACAAAGCTAGTATTTCTGTGAACTCTTAGAGGATT 9479  
Qy 5524 TCGGCTGCTCTCAAAACAACACAGGAGGAAGTAATAAATCATAAATCCCCCATG 5583  
Db 9480 TCGGCTGCTCTCAAAACAACACAGGAGGAAGTAATAAATCATAAAT-CCCCATG 9538  
Qy 5584 GSCCTCCCTTATCATATTTTCTCTKTASTGTTSTTTVACCCSTTTTCACTCTCACTGCA 5643  
Db 9539 GSCCTCCCTTATCATATTTTCTCTTACTGTCTTTTACCCTCTTCACTCTCACTGCA 9598  
Qy 5644 CCCCCTCCATGCGGTATGACACAGTAGTCTCCCTYACCMAGAGTTTCTATGGAGAATG 5703  
Db 9599 CCCCCTCCATGCGGTATGACACAGTAGTCTCCCTTACCAAGAGTTTCTATGGAGAATG 9658  
Qy 5704 CAGCTCCCGGAATATTGATGCCCTGATAGGAGTCTTSTAAGGGAACCCCAACC 5763  
Db 9659 CAGCTCCCGGAATATTGATGCCCTGATAGGAGTCTTCTAAGGGAACCCCAACC 9718  
Qy 5764 TTCACTGCCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCAATGCAT 5823  
Db 9719 TTCACTGCCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCAATGCAT 9778  
Qy 5824 GCAATATCATATTATGGACAGGAAAAATGATTAAATCTCTAGTTGCTCTGGAGGACTTGA 5883  
Db 9779 GCAATATCATATTATGGACAGGAAAAATGATTAAATCTCTAGTTGCTCTGGAGGACTTGA 9838  
Qy 5884 GTCACTGTCTGTGGACTTACTTCAACCAACTGATGCTCTGATGGGGGTGGAGTTCAA 5943  
Db 9839 GTCACTGTCTGTGGACTTACTTCAACCAACTGATGCTCTGATGGGGGTGGAGTTCAA 9898  
Qy 5944 GATCAGGCAAGAGAAAAATGTAAAGAAATGTAATCTCCCAACTCACCGGGGTACATGGC 6003  
Db 9899 GATCAGGCAAGAGAAAAATGTAAAGAAATGTAATCTCCCAACTCACCGGGGTACATGGC 9958  
Qy 6004 ACTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCGTACCCAT 6063  
Db 9959 ACTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCGTACCCAT 10018  
Qy 6064 ACTCGCTGGTAAGCCTATTTAATACCACTTCTAGGCTCCATGAGTCTCTCGGCCCA 6123  
Db 10019 ACTCGCTGGTAAGCCTATTTAATACCACTTCTAGGCTCCATGAGGCTCTCGGCCCA 10078  
Qy 6124 AACCCCTACTAATGTGTGGATATGCTCCCTCGCTGAACTTCARGCCATATGTTTCAATCCCT 6183  
Db 10079 AACCCCTACTAATGTGTGGATATGCTCCCTCGCTGAACTTCARGCCATATGTTTCAATCCCT 10138  
Qy 6184 GTACCTGAAACAATGGAACTTTCAGCACAGAAATAAACAACACTTCCTGTTTGTAGTGA 6243  
Db 10139 GTACCTGAAACAATGGAACTTTCAGCACAGAAATAAACAACACTTCCTGTTTGTAGTGA 10198  
Qy 6244 CCTCTGTTTCCAAATSTGAAATTAACCCATACCTCAAACTCACCTGTGTAAATTTAGC 6303  
Db 10199 CCTCTGTTTCCAAATSTGAAATTAACCCATACCTCAAACTCACCTGTGTAAATTTAGC 10258  
Qy 6304 AATACCTACATACACAACTCCCAATGCTCAGGTGGGTAACTCTCCCAACACAAATA 6363

Db 10259 AATACCTACATACACAACTCCCAATGCTCAGGTGGGTAACTCTCTCCCAACAAATA 10318  
Qy 6364 GTCTGCTTACCTCAGGAATATTTTGTCTGTGCTACCTCAGCTATCGTTGTTTGAAT 6423  
Db 10319 GTCTGCTTACCTCAGGAATATTTTGTCTGTGCTACCTCAGCTATCGTTGTTTGAAT 10378  
Qy 6424 GGCTCTTCAGAAATCTATGTGCTTCTCTCATTTCTTAGTGCCCCVATGRCCTATCTACACT 6483  
Db 10379 GGCTCTTCAGAAATCTATGTGCTTCTCTCATTTCTTAGTGCCCTATGACCATCTACACT 10438  
Qy 6484 GAAACAAGTTTATACAGTTATGTATATCTAAGCCCGCAACAAAGAGTACCCATTCTT 6543  
Db 10439 GAAACAAGTTTATACAGTTATGTATATCTAAGCCCGCAACAAAGAGTACCCATTCTT 10498  
Qy 6544 CTTTGTGTATAGGACAGGAGTCTAGGTGCACTAGGTACTGCGCATTTGGCGGTATCACA 6603  
Db 10499 CTTTGTGTATAGGACAGGAGTCTAGGTGCACTAGGTACTGCGCATTTGGCGGTATCACA 10558  
Qy 6604 ACCTCTACTCAGTTTCTACTACAAACTATCTCAAGAACTAAATGGGACATGGAACGGTCT 6663  
Db 10559 ACCTCTACTCAGTTTCTACTACAAACTATCTCAAGAACTAAATGGGACATGGAACGGTCT 10618  
Qy 6664 GCGGACTCCCTGGTCACTTCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTCAA 10678  
Db 10619 GCGGACTCCCTGGTCACTTCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTCAA 10678  
Qy 6724 AATCGAAGAGCTTTAGACTYCTAAACCTGABAGGGGAACTGTTTATTTTAGGG 6783  
Db 10679 AATCGAAGAGCTTTAGACTYCTAAACCTGABAGGGGAACTGTTTATTTTAGGG 10738  
Qy 6784 GAAAGATCTCTTATTAATGTTAACTCAATCCGGATCGTCACTGAGAAAGTTTAAAGAAAT 6843  
Db 10739 GAAAGATCTCTTATTAATGTTAACTCAATCCGGATCGTCACTGAGAAAGTTTAAAGAAAT 10798  
Qy 6844 CSAGATCGAATPACAACTGAKACAGAGAGCTTCGAAAACACTTGGACCTCGGGGCTCTCT 6903  
Db 10799 CGAGATCGAATPACAACTGAGAGAGAGCTTCGAAAACACTTGGACCTCGGGGCTCTCT 10858  
Qy 6904 AGCBAATGAGTCCCTGGATTTCTCCCTTTTAGGACCTCTAGGACCTATATATTTGCTA 6963  
Db 10859 AGCBAATGAGTCCCTGGATTTCTCCCTTTTAGGACCTCTAGGACCTATATATTTGCTA 10918  
Qy 6964 CTCTCTTTGGACCTCTGATCTTTTACCTCTTGTGTTAACTTTGCTCTTCCAGAAATCGAA 7023  
Db 10919 CTCTCTTTGGACCTCTGATCTTTTAACTCTTGTGTTAACTTTGCTCTTCCAGAAATCGAA 10978  
Qy 7024 CTTGTRAACTACAAATGGAGCCCAAGATGCAAGTCAAGACTAAGATCTACCGCAGACCC 7083  
Db 10979 CTTGTRAACTACAAATGGAGCCCAAGATGCAAGTCAAGACTAAGATCTACCGCAGACCC 11038  
Qy 7084 CTGGACCGGCTGTTAGCCAGATCTGATGTTAATGATCATCAAGGACCCCTCTCTGAG 7143  
Db 11039 CTGGACCGGCTGTTAGCCAGATCTGATGTTAATGATCATCAAGGACCCCTCTCTGAG 11098  
Qy 7144 GAAATCTCAGCTGACAACTCTACTACGCCCAATTCAGCAGGAGCAGTTTAGACGGT 7203  
Db 11099 GAAATCTCAGCTGACAACTCTACTACGCCCAATTCAGCAGGAGCAGTTTAGACGGT 11158  
Qy 7204 SGTGGCCCAACTCCCAACAGCATTAGGTTTTCTGTTGAGATGGGGGACTGAGAGAC 7263  
Db 11159 SGTGGCCCAACTCCCAACAGCATTAGGTTTTCTGTTGAGATGGGGGACTGAGAGAC 11218  
Qy 7264 AGGACTAGCTGATTTCTTAGCTGATTAAGATCCYTAAGCTTAGSTGGGAAGGTGACC 7323  
Db 11219 AGGACTAGCTGATTTCTTAGCTGATTAAGATCCYTAAGCTTAGSTGGGAAGGTGACC 11278  
Qy 7324 ACATCCACTTTAAACAGGGGCTTGGCAACTTAGVTCACTACCTGACCAATCAGAGAGCTC 7383  
Db 11279 ACATCCACTTTAAACAGGGGCTTGGCAACTTAGVTCACTACCTGACCAATCAGAGAGCTC 11338  
Qy 7384 ACTAAATGCTAATTTAGGCAAGAGCAGGAGGTAAAGAAATAGCCCAATCATYTTATTCMTG 7443  
Db 11339 ACTAAATGCTAATTTAGGCAAGAGCAGGAGGTAAAGAAATAGCCCAATCATCTATTGCTG 11398





QY	4385	CGGGAACACAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGAAACCTCAGAGG-CAG	4443
Db	6683	CGGGAACACAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGAAACCTCAGCAGGACAG	6742
QY	4444	TTTTCTCCCTCGGAGCGGTAGCCACTGAAGAGGAGAAATATCTTTTGCTGCAACTAT	4503
Db	6743	TTTTCTCCCTCGGAGCGGTAGCCACTGAAGAGGAGAAATATCTTTTGCTGCAACTAT	6802
QY	4504	CCAATGGAATATCTTAAACCCCTTCATCAAACTTTCACTTAGGCATGCAGACCCA	4563
Db	6803	CCAATGGAATATCTTAAACCCCTTCATCAAACTTTCACTTAGGCATGCAGACCCA	6862
QY	4564	TCARATGGCCAAATCATTTATCTGACACAGCCCTTTCAAACACTATCAAGCABATAKT	4623
Db	6863	TCARATGGCCAAATCATTTATCTGACACAGCCCTTTCAAACACTATCAAGCABATAGT	6922
QY	4624	CAGGGCTGTGAATGTGCCARAAATAATCCCTGCCTTCATCGCCAAAGCTCTTCAGG	4683
Db	6923	CAGGGCTGTGAATGTGCCAGAGAAATAATCCCTGCCTTCATCGCCAAAGCTCTTCAGG	6982
QY	4684	ARACAAARAAACAGGCATTTACCTGARAARACTGGCAACTGATTTTACCCCAAGCCC	4743
Db	6983	AGAACAAAGAACAGGCATTTACCTGAGAAGACTGGCAACTGATTTTACCCCAAGCCC	7042
QY	4744	ARACCTCAGGATTTCAAGTATCTACTAGTCTGGGTARATATTTTACGGGTTGGGCAAG	4803
Db	7043	ARACCTCAGGATTTCAAGTATCTACTAGTCTGGGTARATATTTTACGGGTTGGGCAAG	7102
QY	4804	GCCTTCCCTGTAGGACAGAAAAAGGCCAAGAGGTAAATAAGGCACCTAGTTTCATGAANTA	4863
Db	7103	GCCTTCCCTGTAGGACAGAAAAAGGCCAAGAGGTAAATAAGGCACCTAGTTTCATGAANTA	7162
QY	4864	ATTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGTTTCCAGGCC	4923
Db	7163	ATTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGTTTCCAGGCC	7222
QY	4924	ACAGTAACCCAGGAGTATCCAGGGTTAGGTATACGATATCACTTACACTGGGCTGA	4983
Db	7223	ACAGTAACCCAGGAGTATCCAGGGTTAGGTATACGATATCACTTACACTGGGCTGA	7282
QY	4984	AGGCCACAGTCCTCAGGAGAGTTCGAGAAATGAATGAAYACTCAAGGACATCTAAA	5043
Db	7283	AGGCCACAGTCCTCAGGAGAGTTCGAGAAATGAATGAAYACTCAAGGACATCTAAA	7342
QY	5044	AAGCAACCCAGGAACCCACCTCACATGGCTGTCTGTGCTATAGCCTTAAAAAGA	5103
Db	7343	AAGCAACCCAGGAACCCACCTCACATGGCTGTCTGTGCTATAGCCTTAAAAAGA	7402
QY	5104	ATCTGCAACTTTCCCAAAAGCAGGACTTAGCCCATAGAAATGCTGTATGGAAGGCC	5163
Db	7403	ATCTGCAACTTTCCCAAAAGCAGGACTTAGCCCATAGAAATGCTGTATGGAAGGCC	7462
QY	5164	TTCCATAACCAATGACCTGTGCTGACCCAGACGCCAATAGTTAGTGACATCACT	5223
Db	7463	TTCCATAACCAATGACCTGTGCTGACCCAGACGCCAATAGTTAGTGACATCACT	7522
QY	5224	CCTTAGCCAAATATCAACAGTTCTTAAAAACATTACAGGAACCTATCCTGAGAGAGG	5283
Db	7523	CCTTAGCCAAATATCAACAGTTCTTAAAAACATTACAGGAACCTATCCTGAGAGAGG	7582
QY	5284	GAAGAACAATTATCCACCCWGTGACATGGTATAGTCAAGTCCCTTCYCTCTAATTC	5343
Db	7583	GAAGAACAATTATCCACCCWGTGACATGGTATAGTCAAGTCCCTTCYCTCTAATTC	7642
QY	5344	CATCCCTAGATACATCTCGGAAGACCCCTACCCAGTCAATTTATTTATCCCAATCG	5403
Db	7643	CATCCCTAGATACATCTCGGAAGACCCCTACCCAGTCAATTTATTTATCTACCCCAATCG	7702
QY	5404	TTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGGAGTCAAAATCCTGGACTGC	5463
Db	7703	TTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTGGAGTCAAAATCCTGGACTGC	7762

QY	5464	CAAAGGAACCTGAAATTCAGGAGACAAACGCTAGCTATTCTCTGTGAACCTCTAGAGGATT	5523
Db	7763	CAAAGGAACCTGAAATTCAGGAGACAAACGCTAGCTATTCTCTGTGAACCTCTAGAGGATT	7822
QY	5524	TGCGCTGCTCTTCAAACAACACAGGAGGAAGTAATAAATCATAAATCCCCCATG	5583
Db	7823	TGCGCTGCTCTTCAAACAACACAGGAGGAAGTAATAAATCATAAATCCCCCATG	7881
QY	5584	GSCTCCCTTATCATATTTTCTCTKATSTGTTSTTTTACCCCTSTTTCACTCTCACTGCA	5643
Db	7882	GSCTCCCTTATCATATTTTCTCTKATSTGTTSTTTTACCCCTSTTTCACTCTCACTGCA	7941
QY	5644	CCCCCTCCATGCGCTGTATGACAGTAGCTCCCTTCYACMAGAGTTTCTATGGAGATG	5703
Db	7942	CCCCCTCCATGCGCTGTATGACAGTAGCTCCCTTCYACMAGAGTTTCTATGGAGATG	8001
QY	5704	CAGCGTCCCGAAATATTCATGCGCCCATCGTATAGAGTCTTTTAAAGGAAACCCCAACC	5763
Db	8002	CAGCGTCCCGAAATATTCATGCGCCCATCGTATAGAGTCTTTTAAAGGAAACCCCAACC	8061
QY	5764	TTCACTGCCCCACACCCCATATGCCCGCAACTGCTATCACTCTGCCCCACTCTTTGCATGCAT	5823
Db	8062	TTCACTGCCCCACACCCCATATGCCCGCAACTGCTATCACTCTGCCCCACTCTTTGCATGCAT	8121
QY	5824	GCMAATCACTATTTATGGACAGAAAAATGATTAATCTCTAGTTGCTCTGGAGGACTTGA	5883
Db	8122	GCMAATCACTATTTATGGACAGAAAAATGATTAATCTCTAGTTGCTCTGGAGGACTTGA	8181
QY	5884	GTCACTGCTGTGGACTTACTTCAACCAACTGCTATGCTGATGGGGGTGGAGTTCAA	5943
Db	8182	GTCACTGCTGTGGACTTACTTCAACCAACTGCTATGCTGATGGGGGTGGAGTTCAA	8241
QY	5944	GATCAGGCAAGAGAAAAACATGTAAAGAAATGATCTCCCAACTCACCCGGGTACATGGC	6003
Db	8242	GATCAGGCAAGAGAAAAACATGTAAAGAAATGATCTCCCAACTCACCCGGGTACATGGC	8301
QY	6004	ACCTTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCCAT	6063
Db	8302	ACCTTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCCAT	8361
QY	6064	ACTCGCTGGTAAAGCTATTTAATACCCCTCCTCTGCGCTCCATGAGGTCTCGGCCAA	6123
Db	8362	ACTCGCTGGTAAAGCTATTTAATACCCCTCCTCTGCGCTCCATGAGGTCTCGGCCAA	8421
QY	6124	AAACCTTACTTAACTGTGTGGATATGCTCCCTGGAACCTTCARGCCATATGTTTCAATCCCT	6183
Db	8422	AAACCTTACTTAACTGTGTGGATATGCTCCCTGGAACCTTCARGCCATATGTTTCAATCCCT	8481
QY	6184	GTACCTGAAACAAATGGAAACAACTTCAGCAGAGAAATAAACACACCTTCCTTTAGTAGGA	6243
Db	8482	GTACCTGAAACAAATGGAAACAACTTCAGCAGAGAAATAAACACACCTTCCTTTAGTAGGA	8541
QY	6244	CCTCTGTTTCAATSTGGAATACCCATACCTCAACCTCACCTGTGTAAATTTAGC	6303
Db	8542	CCTCTGTTTCAATSTGGAATACCCATACCTCAACCTCACCTGTGTAAATTTAGC	8601
QY	6304	AATACTACATACACAAACCTCCCAATGCTATGAGTGGTAACTCTCCCAACAAATA	6363
Db	8602	AATACTACATACACAAACCTCCCAATGCTATGAGTGGTAACTCTCCCAACAAATA	8661
QY	6364	GTCTGCTTACCTCAGGAATATTTTGTGTGTGTACCTCAGCCCTATCGTTGTTGAT	6423
Db	8662	GTCTGCTTACCTCAGGAATATTTTGTGTGTGTACCTCAGCCCTATCGTTGTTGAT	8721
QY	6424	GGCTCTTACGAATCTATGTGCTTCTCTCAATCTTTAGTGGCCCCCAATGRCATCTACCT	6483
Db	8722	GGCTCTTACGAATCTATGTGCTTCTCTCAATCTTTAGTGGCCCCCAATGRCATCTACCT	8781
QY	6484	GAACAAGATTTATACAGTTATGCTATCTTAAGCCCGCCCAACAAAGAGTACCCATCTT	6543
Db	8782	GAACAAGATTTATACAGTTATGCTATCTTAAGCCCGCCCAACAAAGAGTACCCATCTT	8841
QY	6544	CCTTTGTTATAGGAGCAGGAGTGTAGGTGCACTAGGTACTGGGCACTTGGCGGTATACA	6603



```
Db      8842  CTTTTTGTATAGGAGCAGAGTGTAGTGCACCTAGGTACTGGCATTTGGCGGTATCACA 8901
QY      6604  ACCTCTACTAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTC 6663
Db      8902  ACCTCTACTAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTC 8961
QY      6664  GCGGACTCCCTGGTCACCTGCAAGATCACTTAACCTCCCTAGCAGCAGTGTCTTCRA 6723
Db      8962  GCGGACTCCCTGGTCACCTGCAAGATCACTTAACCTCCCTAGCAGCAGTGTCTTCRA 9021
QY      6724  AATCGAAGAGCTTTAGACTGCTTAACCGCTGAGAGAGGGGAACTGTTTATTTTAGGG 6783
Db      9022  AATCGAAGAGCTTTAGACTGCTTAACCGCTGAGAGAGGGGAACTGTTTATTTTAGGG 9081
QY      6784  GAAGAATGCTGTATTATGTTAATCAATCCGGAATCGTCACCTGAGAAAGTTTAAAGAAAT 6843
Db      9082  GAAGAATGCTGTATTATGTTAATCAATCCGGAATCGTCACCTGAGAAAGTTTAAAGAAAT 9141
QY      6844  CSAGATCGAATCAACGTTAKAGCAGAGAGCTTCGAAACACTGGACCTGGGGCCCTCCTC 6903
Db      9142  CGAGATCGAATCAACGTTAGCAGAGAGAGCTTCGAAACACTGGACCTGGGGCCCTCCTC 9201
QY      6904  AGCGRATGATGCCCTGGATTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATTGCTA 6963
Db      9202  AGCGRATGATGCCCTGGATTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATTGCTA 9261
QY      6964  CTCCTCTTTGGACCCCTGTATCTTTTRACCTCTTTGTTAACTTTGTCTCTTCAGAAATCGAA 7023
Db      9262  CTCCTCTTTGGACCCCTGTATCTTTTAACTCTCTTTGTTAACTTTGTCTCTTCAGAAATCGAA 9321
QY      7024  GCTGTAAATCAATAATGAGCCCAAGATGCAAGTCAAGACTAAGACTACCGCGAGACC 7083
Db      9322  GCTGTAAATCAATAATGAGCCCAAGATGCAAGTCAAGACTAAGACTACCGCGAGACC 9381
QY      7084  CTGGACCGGCTGTAGCCCAAGATGCTGATGTTTAAATGACATCAAGAGGCAACCTCTGAG 7143
Db      9382  CTGGACCGGCTGTAGCCCAAGATGCTGATGTTTAAATGACATCAAGAGGCAACCTCTGAG 9441
QY      7144  GAAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGGAGCAAGTTAGACGGT 7203
Db      9442  GAAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGGAGCAAGTTAGACGGT 9501
QY      7204  SGTGGCCAACTCCGCCAAGCACTAGCTTTCTGTTGAGATGGGGACTGAGAGAC 7263
Db      9502  C-TGGCCAACTCCGCCAAGCACTAGCTTTCTGTTGAGATGGGGACTGAGAGAC 9560
QY      7264  AGGACTAGCTGGATTTCTTAGCTGATTAAGAAATCCYTAAGCTTAGSTGGGAAGGTGACC 7323
Db      9561  AGGACTAGCTGGATTTCTTAGCTGATTAAGAAATCCYTAAGCTTAGSTGGGAAGGTGACC 9620
QY      7324  ACATCCACTTTTAAACAAGGGGTTGCAACTTAGYTCAACCTGACCAATACAGAGGCTC 7383
Db      9621  ACATCCACTTTTAAACAAGGGGTTGCAACTTAGYTCAACCTGACCAATACAGAGGCTC 9680
QY      7384  ACTAAATGCTAATTAGGCAAGAGAGGATTAAGAAATAGCCATCATYTTATTGCMTG 7443
Db      9681  ACTAAATGCTAATTAGGCAAGAGAGGATTAAGAAATAGCCATCATYTTATTGCMTG 9740
QY      7444  AGAGCACAGCAGAGGAGCAATGATCGGATATAAACCAAGTTTTCGAGCCGGCAACGG 7503
Db      9741  AGAGCACAGCAGAGGAGCAATGATCGGATATAAACCAAGTTTTCGAGCCGGCAACGG 9800
QY      7504  CAACCCCTTTGGGTCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCACCTCT 7563
Db      9801  CAACCCCTTTGGGTCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCACCTCT 9860
QY      7564  ATTAAATCTTGCACTGCR 7582
Db      9861  ATTAAATCTTGCACTGCA 9879
```

```
US-09-573-080A-21
; Sequence 21, Application US/09573080A
; Patent No. 682897
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL
; APPLICANT: JOAN, PETER
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATI
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/573,080A
; CURRENT FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 8523
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat region
; LOCATION: (1)..(8523)
; OTHER INFORMATION: herv17
; PUBLICATION INFORMATION:
; PUBLICATION: Jurka, J; Malchiewicz, J; Milosavljevic, A
; TITLE: Prototypic sequences for human repetitive DNA
; JOURNAL: Journal of Molecular Evolution
; VOLUME: 35
; ISSUE: 4
; PAGES: 286-291
; DATE: 1992-10-
; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
; DATABASE ENTRY DATE:
; DATABASE ENTRY DATE: 1996-01-26
US-09-573-080A-21
```

```
Query Match      81.8%; Score 6124.6; DB 3; Length 8523;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 6363; Conservative 186; Mismatches 50; Indels 53; Gaps 27;

QY      606  GATGGAAACGTTCCCGCAAGACAGAAACGCCCTTAAGACGTATTCTGGAPAAATTGGGA 665
Db      1923  GATGGAAACGTTCCCGCAAGACAGAAACGCCCTTAAGACGTATTCTGGAGAAATTGGGA 1982
QY      666  MCAATTTGACCTCTAGACACTAAGAAAGAAAGACTTATTTCTTCAGTGCCGCTG 725
Db      1983  CCAATTTGACCTCTAGACACTAAGAAAGAAAGACTTATTTCTTCAGTGCCGCTG 2042
QY      726  GCATCTCTGAGGAGGTATAAATTATAACCATCTTACAGCTAGACTCTTTTGTAGAA 785
Db      2043  GCATCTCTGAGGAGGTATAAATTATAACCATCTTACAGCTAGACTCTTTTGTAGAA 2102
QY      786  AAGCAATGAGTGAAGTGCATTAAGTACAACTTTCTTTTTCATTAAGAGACAATCAC 845
Db      2103  AAGCAATGAGTGAAGTGCATTAAGTACAACTTTCTTTTTCATTAAGAGACAATCAC 2162
QY      846  AATTATGTAAAGTGTGATTTATGCTTACAGGAAGCTTTCAGAGTCTACCTCCCTATC 905
Db      2163  AATTATGTAAAGTGTGATTTATGCTTACAGGAAGCTTTCAGAGTCTACCTCCCTATC 2222
QY      906  CCAGCATCCCGACTCTCTTCCCAATTAAGGACCCCTTCAACCCCAATAGTCCAA 965
Db      2223  CCAGCATCCCGACTCTCTTCCCAATTAAGGACCCCTTCAACCCCAATAGTCCAA 2282
QY      966  AAGGAGATAGACAAAGGGTAAACAGTGAAACAAAGAGTGCCAAATTTCCCAATATGA 1025
Db      2283  AAGGAGATAGACAAAGGGTAAACAGTGAAACAAAGAGTGCCAAATTTCCCAATATGA 2342
QY      1026  CCCCTCCCAAGCAGTGAGGAGGAGAAATTCGGCCACGACGAGTGATGTCYTTTYY 1085
Db      2343  CCCCT-CCAGCAGTGAGGAGGAGAAATTCGGCCACGACGAGTGATGTCYTTTTC 2401
QY      1086  TCTCCAGACTTAAAGCAAAATAAAACAGACTTAGGTAAATTTCTCAGATAAATCTGATGG 1145
Db      2402  TCTCCAGACTTAAAGCAAAATAAAACAGACTTAGGTAAATTTCTCAGATAAATCTGATGG 2461
```





Db 4589 CACTAATAAGGAACTCAGAAAGCCATACCCATTTAGTAGATGGACAACTGAAGTAG 4648  
QY 3365 AAGTGGCTTTCCAGGCCCTTAAGAAGCGCTTAAACCCCAAGYCCAGAGTGTAAAGYTTGCC 3424  
Db 4649 AAGTGGCTTTCCAGGCC-----TAAACCCCAAGCCCAAGTGTAAAGTGTGCC 4694  
QY 3425 AACGGGCAAGACTTTTSTTVATAYRTACACAGAAAAACAGAAAYAGCTCTGCGAGTCC 3484  
Db 4695 AACAGGGCAAGACTTTTTCATATGTACAG-AAAAAACAGGAATAGCTCTAGGAGTCC 4753  
QY 3485 TTACACAGTCCRAGGGAYGAGCTTGCACCCYRTGTCRVAACCTGASTAAGGAAAAVTGATG 3544  
Db 4754 TTACACAGATCCGAGGGATGACTTGCACCTGTGGCATACTGACTAAGGAAATGATG 4813  
QY 3545 TAGTGGCAAGGGTGTGRCYTCATTTGTAYGGGTAGTGGTGGCAGTAGCAGYKTAGTAT 3604  
Db 4814 TAGTGGCAAGGGTGTGACCTCATTTGTTACGGGTAGTGGTGGCAGTAGCAGTCTTAGTAT 4873  
QY 3605 CTGAAGCAGTTAAATATAACAGGGBAGAGATCTTACTGTGTGGACATCTCATGAKGTGA 3664  
Db 4874 CTGAAGCAGTTAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTA 4933  
QY 3665 AYRGCATACTCACTGCTAAAGGAGACTTGTGGCTGTGACACAACYGTTTACTTAAATRTC 3724  
Db 4934 ATGGCATACTCACTGCTAAAGGAGACTTGTGGCTGTGACACAACYGTTTACTTAAATGTC 4993  
QY 3725 AGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACTTGTGGCACTCTTAAACCCAG 3784  
Db 4994 AGGCTCTATTACTTGAAGGGCCAGTGTGCGACTGTGCACTTGTGCAACTCTTAAACCCAG 5053  
QY 3785 YCNCAATTTCTCCAGACATGAGAAAGATARAAYATACTGTCAAACAAATTAATTTCTC 3844  
Db 5054 CCAACATTTCTCCAGACATGAGAAAGATARAACATAAACAATAACTGTCAACAAATTAATTTCTC 5113  
QY 3845 AAACCTATGCCACTCGAGGGGACCTTGTAGAGTTCCTTGTACTGATCCYGCACCTTCAAC 3904  
Db 5114 AAACCTATGCCACTCGAGGGGACCTTTTAGAGTTCCTTTGACTGATCCCGACC-TCAAC 5172  
QY 3905 TTGTATACGTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAAGYGGGGTATGAGTG 3964  
Db 5173 TTGTATACGTGATGGAAGTTCCTTTGTAGAAAAAGGACTTCGAAAAAGTGGGGTATGAGTG 5232  
QY 3965 GTCAGTGATTAATGGAATATTTGAAGTATCCCTCACTCCAGGAACCTAGTCTTAGCTR 4024  
Db 5233 GTCAGTGATTAATGGAATACTTGAAGTATCCCTCACTCCAGGAACCTAGTCTTAGCTR 5292  
QY 4025 GCAGAACTAATAGCCVTCAYTKGGGCACCTAGAAATTAGGAGAGAAAAAGGGYAAATATA 4084  
Db 5293 GCAGAACTAATAGCCCTCACTTGGGCACTAGAAATTAGGAGAGAAAAAGGGCAATATA 5352  
QY 4085 TATACAGACTCTAATATGCTTACCTAGTTCNTCCATGCCCATGMRCAATATGSRAGAA 4144  
Db 5353 TATACAGACTCTAATATGCTTACCTAGTTCNTCCATGCCCATGMRCAATATGSRAGAA 5412  
QY 4145 AGGGAATTCCTAATCTTCYAGRGGAACACCTATCMAACATCAGGAAGCCATTAGGABATTA 4204  
Db 5413 AGGGAATTCCTAATCTTCYAGRGGAACACCTATCMAACATCAGGAAGCCATTAGGABATTA 5472  
QY 4205 TTATGCGWGTACAGAAACCTARAGAGTGGWAGTCTTACACTGCGYGGGTCAATCANAAA 4264  
Db 5473 TTATGCGGTGACAGAAACCTAAGAGGTGGGAGTCTTACACTGCGCGGGGTCAATCANAAA 5532  
QY 4265 GGAAGGAAGGGAATASAAAGGAAATGCGCAAGCAKATATTTGAAGCMAAAAGAGCTGCA 4324  
Db 5533 GGAAGGAAGGGAATASAAAGGAAATGCGCAAGCAKATATTTGAAGCMAAAAGAGCTGCA 5592  
QY 4325 AGGCAGGACCTCCATTAGAATGCTTATTAACCTTCCCTTAGTATAGGGTAAATCCCTTC 4384  
Db 5593 AGGCAGGACCTCCATTAGAATGCTTATTAACCTTCCCTTAGTATAGGGTAAATCCCTTC 5652  
QY 4385 CGGGAACCAAGCCCAAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCAGCAGG-CAG 4443

Db 5653 CGGGAACCAAGCCCAAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCAGCAGGACAG 5712  
QY 4444 TTTTCTCCCTCGGAGCGGTTAGCCACTGAGAGAGGGAATACTTTTTCCTGCAACTAT 4503  
Db 5713 TTTTCTCCCTCGGAGCGGTTAGCCACTGAGAGAGGGAATACTTTTTCCTGCAACTAT 5772  
QY 4504 CCAATGGAAATTAATTAATAACCTTCAATAACCTTTCACCTTAGGCATCGATAGCACCA 4563  
Db 5773 CCAATGGAAATTAATTAATAACCTTTCACAACTTTCACCTTAGGCATCGATAGCACCA 5832  
QY 4564 TCARATGGCCAAATCAATTAATTAATTCGACAGGCGCTTTTCAAACTATCAAGCABATAT 4623  
Db 5833 TCAGATGGCCAAATCAATTAATTAATTCGACAGGCGCTTTTCAAACTATCAAGCABATAT 5892  
QY 4624 CAGGCGCTGTGAATGTGCGARABAAATAATCCCTGCTCTATCCGCAAGCTCTTTCAGG 4683  
Db 5893 CAGGCGCTGTGAATGTGCGARABAAATAATCCCTGCTCTATCCGCAAGCTCTTTCAGG 5952  
QY 4684 ARAACAAARAACAGGCCATTAACCTGARAARAACCTGGCAACTGATTTTACCCACAAGGCC 4743  
Db 5953 ARAACAAARAACAGGCCATTAACCTGARAARAACCTGGCAACTGATTTTACCCACAAGGCC 6012  
QY 4744 AAACCTCAGGGAATTCAGTATCTACTAGTCTGGGTARATACTTTTCAACGGGTGGCCARAG 4803  
Db 6013 AAACCTCAGGGAATTCAGTATCTACTAGTCTGGGTARATACTTTTCAACGGGTGGCCARAG 6072  
QY 4804 GCCTTCCCTCTAGGACAGAAAGGCCCAAGGATTAATAAAGGCATAGTTTCAATGAATA 4863  
Db 6073 GCCTTCCCTCTAGGACAGAAAGGCCCAAGGATTAATAAAGGCATAGTTTCAATGAATA 6132  
QY 4864 ATTCGCCGATTCGCACTTCCCGAGGCTTACAGAGTGAACAATAGCCCTGCTTTCCAGGCC 4923  
Db 6133 ATTCGCCGATTCGCACTTCCCGAGGCTTACAGAGTGAACAATAGCCCTGCTTTCCAGGCC 6192  
QY 4924 ACAGTAAACCCAGGGAGTATCCAGCGTGGTATACGATATCACTTACACTGCGCTGA 4983  
Db 6193 ACAGTAAACCCAGGGAGTATCCAGCGTGGTATACGATATCACTTACACTGCGCTGA 6252  
QY 4984 AGGCCACAGTCTCAGGGAGGTCGAGAAATGAATGAANAATCACTCAAGAGGACATCTAAA 5043  
Db 6253 AGGCCACAGTCTCAGGGAGGTCGAGAAATGAATGAANAATCACTCAAGAGGACATCTAAA 6312  
QY 5044 AAGCAAAACCCAGGAAACCCACTCACATGGCTGTCTGTTGCTCTATAGCCCTTAAAAAGA 5103  
Db 6313 AAGCAAAACCCAGGAAACCCACTCACATGGCTGTCTGTTGCTCTATAGCCCTTAAAAAGA 6372  
QY 5104 ATCTGCAACTTTCCCAAAAAAGCAGGACTTAGCCCATACGAATCTGTATGGAAGGCC 5163  
Db 6373 ATCTGCAACTTTCCCAAAAAAGCAGGACTTAGCCCATACGAATCTGTATGGAAGGCC 6432  
QY 5164 TTCAATAACCAATGACCTTGTGCTTGACCCCAAGACCAACTTAGTTGGCAGACATCACT 5223  
Db 6433 TTCAATAACCAATGACCTTGTGCTTGACCCCAAGACCAACTTAGTTGGCAGACATCACT 6492  
QY 5224 CCTTAGCCAAATATCAACAAGTTCCTTAAAAATTAACAAGAACCTTATCCCTCAGAGAGG 5283  
Db 6493 CCTTAGCCAAATATCAACAAGTTCCTTAAAAATTAACAAGAACCTTATCCCTCAGAGAGG 6552  
QY 5284 GAAAAAGAACTATTCACCCWMTGTCATGTTAGTCAAGTCCCTTCYCTCTAATTC 5343  
Db 6553 GAAAAAGAACTATTCACCCWMTGTCATGTTAGTCAAGTCCCTTCYCTCTAATTC 6612  
QY 5344 CATCCTTAGATACATCTCTGGGAGGACCTTACCCAGATCACTTATTTATTTAGCCCACTGGCG 5403  
Db 6613 CATCCTTAGATACATCTCTGGGAGGACCTTACCCAGATCACTTATTTATTTAGCCCACTGGCG 6672  
QY 5404 TTAAAGTGCCTGGAGTGGAGTCTTCGATACATCACTTTAGTCAAACTCCTGATCTGC 5463  
Db 6673 TTAAAGTGCCTGGAGTGGAGTCTTCGATACATCACTTTAGTCAAACTCCTGATCTGC 6732  
QY 5464 CAAAGGAACCTGAAATCCAGGAGCAACGCTAGCTATTTCTGTGAACTCTTAGAGGAT 5523  
Db 6733 CAAAGGAACCTGAAATCCAGGAGCAACGCTAGCTATTTCTGTGAACTCTTAGAGGAT 6792

QY 5524 TGGCGCTGCTCTTCAACACACACACAGGAGAAAGTAACTAAATCATAAATCCCCCATG 5583  
 DB 6793 TGGCGCTGCTCTTCAACACACACACAGGAGAAAGTAACTAAATCATAAATCCCCCATG 6851  
 QY 5584 GSCCTCCCTTATCATATTTTCTCTASTAGTTSTTTTACCTCTTTTACCTCTCTACCTGCA 5643  
 DB 6852 GSCCTCCCTTATCATATTTTCTCTASTAGTTSTTTTACCTCTTTTACCTCTCTACCTGCA 6911  
 QY 5644 CCCCCTCCATGCGCTGTATGACAGTAGCTCCCTTACCMAGAGTTTCTATGGAGATG 5703  
 DB 6912 CCCCCTCCATGCGCTGTATGACAGTAGCTCCCTTACCMAGAGTTTCTATGGAGATG 6971  
 QY 5704 CAGCTCCCGGAAATATGATGCCCCCATCTATAGAGTCTTTSTAAGGAAACCCCAACC 5763  
 DB 6972 CAGCTCCCGGAAATATGATGCCCCCATCTATAGAGTCTTTSTAAGGAAACCCCAACC 7031  
 QY 5764 TTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTTGCATGCAT 5823  
 DB 7032 TTCACTGCCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTTGCATGCAT 7091  
 QY 5824 GCAATATCTATTTGGACAGGAAATAATGATTAATCTTCTAGTTGTCTTGAGGACTTGG 5883  
 DB 7092 GCAATATCTATTTGGACAGGAAATAATGATTAATCTTCTAGTTGTCTTGAGGACTTGG 7151  
 QY 5884 GTCACTGTCTGTGGACTTACTTCAACCAAACTGGTATGCTGATGGGGTGGAGTTCAA 5943  
 DB 7152 GTCACTGTCTGTGGACTTACTTCAACCAAACTGGTATGCTGATGGGGTGGAGTTCAA 7211  
 QY 5944 GATCAGGAGAGGAAACATGTAAAGAGTAACTCTCCCAACTCAACCGGGTACATGGC 6003  
 DB 7212 GATCAGGAGAGGAAACATGTAAAGAGTAACTCTCCCAACTCAACCGGGTACATGGC 7271  
 QY 6004 ACCTCTAGCCCTTACAAAGSACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCCAT 6063  
 DB 7272 ACCTCTAGCCCTTACAAAGSACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCCAT 7331  
 QY 6064 ACTGCGCTGGTAACTTATTAATACACCTCACTGGGTTCATGAGTCTCGGCGCAA 6123  
 DB 7332 ACTGCGCTGGTAACTTATTAATACACCTCACTGGGTTCATGAGTCTCGGCGCAA 7391  
 QY 6124 AACCTTACTAACTGTGTGGATATGCTCCCTGAACTTCAGCGCATATGTTTCAATCCCT 6183  
 DB 7392 AACCTTACTAACTGTGTGGATATGCTCCCTGAACTTCAGCGCATATGTTTCAATCCCT 7451  
 QY 6184 GTACCTGAACAAATGGAACTTTCAGCAGAGAAATAAACACCACTTCCGTTTTAGTAGCA 6243  
 DB 7452 GTACCTGAACAAATGGAACTTTCAGCAGAGAAATAAACACCACTTCCGTTTTAGTAGCA 7511  
 QY 6244 CTTCTGTGTTCCATSTGGAAATAACCCATACCTCAAACTCACTGTGTAAATTTAGC 6303  
 DB 7512 CTTCTGTGTTCCATSTGGAAATAACCCATACCTCAAACTCACTGTGTAAATTTAGC 7571  
 QY 6304 AATACTACATACACACCACTCCCAATGCATCAGTGGTAACTCTCTCCACACAAATA 6363  
 DB 7572 AATACTACATACACACCACTCCCAATGCATCAGTGGTAACTCTCTCCACACAAATA 7631  
 QY 6364 GTCTGCTACCTCAGGAATATTTTGTGTGGTACCTCAGCCATGCTGTTTGAAT 6423  
 DB 7632 GTCTGCTACCTCAGGAATATTTTGTGTGGTACCTCAGCCATGCTGTTTGAAT 7691  
 QY 6424 GGCTCTTCAAGATCTATGCTCTCTCAATTTTAGTGCCTCCCTCCCTATGCTACACT 6483  
 DB 7692 GGCTCTTCAAGATCTATGCTCTCTCAATTTTAGTGCCTCCCTATGCTACACT 7751  
 QY 6484 GAACAGATTTATAGATTTATGCTATCTAGCCCGGCAACAAAGAGTACCCATCTT 6543  
 DB 7752 GAACAGATTTATAGATTTATGCTATCTAGCCCGGCAACAAAGAGTACCCATCTT 7811  
 QY 6544 CTTTTGTTATAGGACGAGTGTGCTAGTGGCACTAGGTACCTGGCATTCGCGGTATCACA 6603  
 DB 7812 CTTTTGTTATAGGACGAGTGTGCTAGTGGCACTAGGTACCTGGCATTCGCGGTATCACA 7871

QY 6604 ACCTCTACCTCAGTTCTACTACAAACTATCTCAAGAACTATAAATGGGACATGGAAACGGGTC 6663  
 DB 7872 ACCTCTACCTCAGTTCTACTACAAACTATCTCAAGAACTATAAATGGGACATGGAAACGGGTC 7931  
 QY 6664 GCGGACTCCCTGCTGCTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTCRA 6723  
 DB 7932 GCGGACTCCCTGCTGCTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTCRA 7991  
 QY 6724 AATCGAAGAGCTTTAGACTYGTCTAACCGCTGAPAGAGGGGAAACCTGTTTATTTTAGGG 6783  
 DB 7992 AATCGAAGAGCTTTAGACTYGTCTAACCGCTGAPAGAGGGGAAACCTGTTTATTTTAGGG 8051  
 QY 6784 GAGAGTCTGTTTATTTATTTATCTTAATCGGATCTCTCACTGAGAAAGTTTAAAGAAAT 6843  
 DB 8052 GAGAGTCTGTTTATTTATTTATCTTAATCGGATCTCTCACTGAGAAAGTTTAAAGAAAT 8111  
 QY 6844 CSAGATCGAATCAACAGTAKAGCAGAGAGCTTCGAAACACTGAGCCCTGGGSCCTCCTC 6903  
 DB 8112 CGAGATCGAATCAACAGTAKAGCAGAGAGCTTCGAAACACTGAGCCCTGGGSCCTCCTC 8171  
 QY 6904 AGCCATGGATGCCCTGGATTTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATATTTGCTA 6963  
 DB 8172 AGCCATGGATGCCCTGGATTTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATTTGCTA 8231  
 QY 6964 CTCTCTTTGGACCTCTGATCTTTTACCTCTTTTAACTTTTCTCTCCAGAAATCGAA 7023  
 DB 8232 CTCTCTTTGGACCTCTGATCTTTTAACTCTTTTAACTTTTCTCTCCAGAAATCGAA 8291  
 QY 7024 GCTGTAAACTCAAAATGGAGCCCAAGATGCACTCCAGACTAAGTCTACCGCAGACCC 7083  
 DB 8292 GCTGTAAACTCAAAATGGAGCCCAAGATGCACTCCAGACTAAGTCTACCGCAGACCC 8351  
 QY 7084 CTGGACCGGCTCTGTAGCCCAAGATCTGATGTTTAACTCAATCAAGGCAACCTCTCTGAG 7143  
 DB 8352 CTGGACCGGCTCTGTAGCCCAAGATCTGATGTTTAACTCAATCAAGGCAACCTCTCTGAG 8411  
 QY 7144 GAAATCTCAGTGTGCAACCTCTAATGCCCCCAATTCAGCAGGAAAGCAGTTAGACGGT 7203  
 DB 8412 GAAATCTCAGTGTGCAACCTCTAATGCCCCCAATTCAGCAGGAAAGCAGTTAGACGGT 8471  
 QY 7204 SCTGCGGCAACCTCCCAACAGCAGCTTAGTCTCTGTTGAGTGGGGAC 7255  
 DB 8472 CCGTGGGCAACCTCCCAACAGCAGCTTAGTCTCTGTTGAGTGGGGAC 8523

RESULT 4

US-09-854-867-21  
 ; Sequence 21, Application US/09854867  
 ; Patent No. 7014997  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JOAN, KNOLL H  
 ; APPLICANT: ROGAN, PETER K  
 ; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING  
 ; FILE REFERENCE: 30307  
 ; CURRENT APPLICATION NUMBER: US/09/854,867  
 ; CURRENT FILING DATE: 2003-05-08  
 ; NUMBER OF SEQ ID NOS: 613  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 21  
 ; LENGTH: 8523  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: repeat\_region  
 ; LOCATION: (1)..(8523)  
 ; OTHER INFORMATION: herv17  
 US-09-854-867-21

Query Match 81.8%; Score 6124.6; DB 5; Length 8523;

Best Local Similarity 95.7%; Pred. No. 0;

Matches 6363; Conservative 186; Mismatches 50; Indels 53; Gaps 27;

QY 606 GATGGGAAACGTTCCCGCAAGACAAAGGCGCCCTAAGACGTTATTTCTGGAATTTGGGA 665







|||||  
6253 AGGCCACAGTCTCAGGAGAGGTCAAGAAATGAATGAACACTCAAGGACATCTAATA 6312  
QY  
5044 AAGCAAAACCCAGAAACCCACTCACTGCGCTGVTCTGTGTGCTTATAGCCTTAAAAAGA 5103  
Db  
6313 AAGCAAAACCCAGAAACCCACTCACTGCGCTGCTCTGTGTGCTTATAGCCTTAAAAAGA 6372  
QY  
5104 ATCTGCAACTTTCCTCCAAAGCAGGACTTAGCCCATACGAATGCTGTATGGAAGGCC 5163  
Db  
6373 ATCTGCAACTTTCCTCCAAAGCAGGACTTAGCCCATACGAATGCTGTATGGAAGGCC 6432  
QY  
5164 TTCTAAACCAATGACTTGTGCTTGACCCCAAGACCACTTGTGTGAGACATCACT 5223  
Db  
6433 TTCTAAACCAATGACTTGTGCTTGACCCCAAGACCACTTGTGTGAGACATCACT 6492  
QY  
5224 CTTTAGCCAAATPATCAACAGTTCTTAAACATTAACAAGAACCTTCCCTGAGAGAGG 5283  
Db  
6493 CTTTAGCCAAATPATCAACAGTTCTTAAACATTAACAAGAACCTTCCCTGAGAGAGG 6552  
QY  
5284 GAAAGAACTATTCCACCCWGTGACATGGTATTAGTCAAGTCCCTTCVCTCTAATTCCC 5343  
Db  
6553 GAAAGAACTATTTCACCCCTTGTGACATGGTATTAGTCAAGTCCCTTCVCTCTAATTCCC 6612  
QY  
5344 CATCCTTAGATACATCCTGGGAAGACCTACCCAGTCAATTTATVATACCCCACTGCGG 5403  
Db  
6613 CATCCTTAGATACATCCTGGGAAGACCTACCCAGTCAATTTATVATACCCCACTGCGG 6672  
QY  
5404 TTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTTGAGTCAAACTCTGGATCTGC 5463  
Db  
6673 TTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTTGAGTCAAACTCTGGATCTGC 6732  
QY  
5464 CAAAGAACTGAAATCAGAGAGACCAAGCTAGCTATTCCTGTGAACTCTTAGAGGATT 5523  
Db  
6733 CAAAGAACTGAAATCAGAGAGACCAAGCTAGCTATTCCTGTGAACTCTTAGAGGATT 6792  
QY  
5524 TCGCGCTGCTCTTCAAAACCAACAGAGGAGGAACTAACTAAATCATAAATCCCCCATG 5583  
Db  
6793 TCGCGCTGCTCTTCAAAACCAACAGAGGAGGAACTAACTAAATCATAAAT - CCCCATG 6851  
QY  
5584 GSCCTCCCTTATCATATTTTCTCTKTASTGTSTTTTACCCTSTTTCACTCTCACTGCA 5643  
Db  
6852 GSCCTCCCTTATCATATTTTCTCTTACTGTTCTTTTACCCTCTTCACTCTCACTGCA 6911  
QY  
5644 CCCCCTCGATGCGGTATGATACCGTAGCTCCCTYACCMAGAGTTTCTATGGAGATG 5703  
Db  
6912 CCCCCTCGATGCGGTATGATACCGTAGCTCCCTTACCAAGAGTTTCTATGGAGATG 6971  
QY  
5704 CAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCAAC 5763  
Db  
6972 CAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCAAC 7031  
QY  
5764 TTCACTGCCCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCAATGCAT 5823  
Db  
7032 TTCACTGCCCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCAATGCAT 7091  
QY  
5824 GCAAAATACATTAATTGGACAGGAAATAATGATTAATCCCTAGTTGCTCTGGAGGACTTGA 5883  
Db  
7092 GCAAAATACATTAATTGGACAGGAAATAATGATTAATCCCTAGTTGCTCTGGAGGACTTGA 7151  
QY  
5884 GTCACTGCTGTGGACTTACTTCAACCAACTGGTATGCTGTAGTGGGGGTGGAGTTCAA 5943  
Db  
7152 GTCACTGCTGTGGACTTACTTCAACCAACTGGTATGCTGTAGTGGGGGTGGAGTTCAA 7211  
QY  
5944 GATCAGGCAAGGAAACCATGTAAAGAAAGTAATCTCCCAACTCAACCGGGGTACATGGC 6003  
Db  
7212 GATCAGGCAAGGAAACCATGTAAAGAAAGTAATCTCCCAACTCAACCGGGGTACATGGC 7271  
QY  
6004 ACTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCCAT 6063  
Db  
7272 ACTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACCCAT 7331  
QY  
6064 ACTCGCTCGTAAAGCTATTTAATACCACCTCACCTGGGCTCCATGAGGTCTCGGCCAA 6123  
|||||

Db  
7332 ACTCGCTCGTAAAGCTATTTAATACCACTCACTGGGCTCCATGAGGTCTCGGCCCAA 7391  
QY  
6124 AACCTTACTAATGTTGGATATGCTCCCTCCCTGAACCTTCARGCCATATGTTTCAATCCCT 6183  
Db  
7392 AACCTTACTAATGTTGGATATGCTCCCTCCCTGAACCTTCARGCCATATGTTTCAATCCCT 7451  
QY  
6184 GTACTCTGAACAACTGGAACAACTTCAGCACAGAAATAAACACACCTTCGTTTCTAGTAGA 6243  
Db  
7452 GTACTCTGAACAACTGGAACAACTTCAGCACAGAAATAAACACACCTTCGTTTCTAGTAGA 7511  
QY  
6244 CCTCTGTTTCCAAATSTGGAATAAACCCATACCTCAAACTCACCTGTGTGTAAATTTAGC 6303  
Db  
7512 CCTCTGTTTCCAAATSTGGAATAAACCCATACCTCAAACTCACCTGTGTGTAAATTTAGC 7571  
QY  
6304 AATACTACATACAAACCACTCCCAATGCAATCAGGTGGGTAACTCTCTCCCAACAAATA 6363  
Db  
7572 AATACTACATACAAACCACTCCCAATGCAATCAGGTGGGTAACTCTCTCCCAACAAATA 7631  
QY  
6364 GTCTGCCCTPACCTCAGGAATAATTTTGTCTGTGTACTCTCAGCCCTATCGTTGTTGAAT 6423  
Db  
7632 GTCTGCCCTPACCTCAGGAATAATTTTGTCTGTGTACTCTCAGCCCTATCGTTGTTGAAT 7691  
QY  
6424 GGCTCTTCAGAACTCTATGTGCTTCTCTCATCTTAGTGCCCCCVATGRCACTTACACT 6483  
Db  
7692 GGCTCTTCAGAACTCTATGTGCTTCTCTCATCTTAGTGCCCCCTATGACCATCTACACT 7751  
QY  
6484 GAAACAAGATTTATACAGTTATGTCAATCTAAGCCCGCAACAAAGAAGTACCCATCTT 6543  
Db  
7752 GAAACAAGATTTATACAGTTATGTCAATCTAAGCCCGCAACAAAGAAGTACCCATCTT 7811  
QY  
6544 CCTTTGTTATAGGAGAGGAGTGTAGTGCACCTAGGTACTTGGCATTTGGCGGTATCACA 6603  
Db  
7812 CCTTTGTTATAGGAGAGGAGTGTAGTGCACCTAGGTACTTGGCATTTGGCGGTATCACA 7871  
QY  
6604 ACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGAACATGGAAACGGGTC 6663  
Db  
7872 ACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGGAACATGGAAACGGGTC 7931  
QY  
6664 GCCGACTCCCTGGTCACTTGCAGAGATCAATTTAACTCCCTAGCAGCAGTAGTCTTCBA 6723  
Db  
7932 GCCGACTCCCTGGTCACTTGCAGAGATCAATTTAACTCCCTAGCAGCAGTAGTCTTCBA 7991  
QY  
6724 AATCGAAGAGCTTTAGACTGTCTAACCGCTGACAGAGGGGAACCTGTTTATTTTAGGG 6783  
Db  
7992 AATCGAAGAGCTTTAGACTGTCTAACCGCTGACAGAGGGGAACCTGTTTATTTTAGGG 8051  
QY  
6784 GAAGAACTCTGTATTATTATGTTAACTCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAAT 6843  
Db  
8052 GAAGAACTCTGTATTATTATGTTAACTCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAAT 8111  
QY  
6844 CSAGATCGAATACAAACGTAAGAGAGCTTTGAAACATCTGGAACCTCTGGGGCTCTCCTC 6903  
Db  
8112 CGAGATCGAATACAAACGTAAGAGAGAGCTTTGAAACATCTGGAACCTCTGGGGCTCTCCTC 8171  
QY  
6904 AGCCBATGGATGCTGGAATCTCCCTTCTTAGGACCTCTAGCAGCTATATATTTGCTA 6963  
Db  
8172 AGCCBAATGGATGCTGGAATCTCCCTTCTTAGGACCTCTAGCAGCTATATATTTGCTA 8231  
QY  
6964 CTCTCTTTGGAACCTGTATCTTTTACCTCTCTTGTAACTTTTGTCTCTTCCAGAAATCGAA 7023  
Db  
8232 CTCTCTTTGGAACCTGTATCTTTTAACTCTCTTGTAACTTTTGTCTCTTCCAGAAATCGAA 8291  
QY  
7024 GCTGTAAACTACAAATGGAGCCCAAGATGAGTCCAAAGATCTAAGATCTACCGAGACCC 7083  
Db  
8292 GCTGTAAACTACAAATGGAGCCCAAGATGAGTCCAAAGATCTAAGATCTACCGAGACCC 8351  
QY  
7084 CTGGACCGGCTGTGTAGCCCGATCTGATGTTTAAATGATCATCAAGGACACCCCTCTGAG 7143  
Db  
8352 CTGGACCGGCTGTGTAGCCCGATCTGATGTTTAAATGATCATCAAGGACACCCCTCTGAG 8411  
QY  
7144 GAAATCTCAGGTGCAACACTCTACTAGCCCCCAATTCAGCAGGAGGAGCTTAGAGCGGT 7203  
Db  
8412 GAAATCTCAGGTGCAACACTCTACTAGCCCCCAATTCAGCAGGAGGAGCTTAGAGCGGT 8471







```
Db 16205 --ATGGCCCCCGGCGTCATATTTTCTCTTACTGTGTCTTACCCCGCTTTCACATATCA 16148
Qy 5639 CTGACCCCTCCATGCGCTGTATGACCACTAGTCCCTTACCMAGAGTTTCTATGGA 5698
Db 16147 CCTCACCCCTCCATGCGCTATACCACTAGTCCCTTACCAAGAGCTTCTATGGA 16088
Qy 5699 GAATGAGCGCTCCCGGAATATTGATGCCCGCCATCGTATAGGAGTCTTTTAAAGGGAACCC 5758
Db 16087 GAATGCGGCTTCCAGAAATATTGATGCCCGCCATCGTATAGGAGTCTTTTCTAAAGGAACC 16028
Qy 5759 CCACCTTCACTGCCCAACCCATATGCCCGCAACTGC 5796
Db 16027 CCACCTTCACTGCCCAACCCATATGCCCGCTGCATTC 15990

RESULT 6
US-09-949-016-12249/c
; Sequence 12249, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12249
; LENGTH: 77997
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(77997)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12249

Query Match 40.0%; Score 2994.2; DB 3; Length 77997;
Best Local Similarity 85.5%; Pred. No. 0;
Matches 3418; Conservative 149; Mismatches 362; Indels 69; Gaps 23;

Qy 1812 ACTTTGGAGGCTCTGGAAAGGGAAAGCTGGGCAAAATTGAATGCCCTAATAGGGCTTGCT 1871
Db 19935 ACTTTGGAGGCTCTGGAAAGGGAAAGCTGGGCAAAATCAATGCCCTGATAGGGCTTGCT 19876

Qy 1872 TCCAGTGGCGGTCTAGAAGACACTTTAAAGAGATTTGTCCAGTAGAAGTAAGCCGCCCC 1931
Db 19875 TCCAGTGGCGGTCTAGAAGACACTTTAAAGAGATTTGTCTGAATAGAAATAAGCTGCCCC 19816

Qy 1932 TTCTGTCATGCCCTTTATTTCAAGGGAATCACTTGAAGGCCCACTGCCCGGAGGACAAA 1991
Db 19815 -TCATCCATGCCCTCATGTCAAGGGAATCACTTGAAGGCCCACTGCCCGGAGGATGAA 19757

Qy 1992 GGTCTTTTGGTTCAGAGCCACTTAACAGATGATCCAGCAGCAGGACTGAGGGTGCTGG 2051
Db 19756 GGTCTTTTGGTTCAGAGCCACTTAACAGATGATCCAGCAGCAGGACTGAGGGTGCTGG 19697

Qy 2052 GGCAGGGCCATCCCATGCCATCAACCTCAGAGGCCCTGGGTATGCTTGAACATGAGG 2111
Db 19696 GGCAGGGCCAGGCCAT-----GCCCGGAGGTATGCTTGAACATGAGG 19654

Qy 2112 GCCAGGAAGT---TCTCTCCTGGACACTGGTGGGTCTTCTTAGTCTTACTTCTTCTGTC 2168
Db 19653 GCCAGGAAGGTAACTGCTCTCCTGAACTGGCATGGCCCTTCTCAGTCTTAGTCTCTGGTCT 19594
```

```
Qy 2169 CCGGACAACTGTCTCTCCAGATCTGTCACTATTCTGAGGGGTCCTNTAAGACGGGAGTCA 2228
Db 19593 CTGGACAACCGTCTCTCCAGATCTGTCACTACCC---GAGGGGTCCTAGGATAGGCAATCA 19537

Qy 2229 CTAGATACTTTTCCAGCCACTTAAGTATGAATCGGAGCTTTATCTCTTTTCAATGC 2288
Db 19536 CTAGATAC-TTCTCCAGCCACTTAAGTTGTG-ACTGGGAACTTCACTCTTTTCACTTGC 19479

Qy 2289 TTTTCTAAATATGCTTTGAAAGCCCACTACCTTTGTTAGGGAGAGACATTTCTAGCAAAAGC 2348
Db 19478 CTTTCTAAATATGCTTTGAAAGCCCACTACCTTTGTTAGGGAGATACATTTCTAGCAAAAGC 19419

Qy 2349 AGGGGCCATATACACTGAACATAGGAGAGAAACACCGCTTTGTTGTTGTTGTTGTTGTTG 2408
Db 19418 AGGGGCCATATATACACTGAACATAGGAGAGAAACACCGTTTGTCT--GTCCCTACTTGT 19361

Qy 2409 AGGAAGGAATTAATCTGAAGTCTGGGCAACAGAGGACAAATATGACGACGAGCCAAAGAAT 2468
Db 19360 AGGAAGGAATTAATCTGAAGTCTGGGCAACAGAGGACAAATATGATGAG-GGAAGAAT 19302

Qy 2469 GCCGCTCTGTTCAAGTTAAACTTAAAGATTCACCTTCTCTTCCCTTACCAAGGCGAGTAC 2528
Db 19301 GCCGCTCTGTTCAAGTTAAACTTAAAGATTCGCGCTCTTCTCCCAACCAAGGCGAGTAC 19242

Qy 2529 CCCTCAGACCCCAAGGCCCAACAGGATTCGAAAGATTTGTTAAGGACTTAAAGCCCAA 2588
Db 19241 CCTCTTAGACCCGAGGCCCAACAGGATTCGAAAGATTTGTTAAGGACCTTAAAGCCCAA 19182

Qy 2589 GCTTAGTAAACCACTGCACTACTCCCTGACGATTAATCCGTAGTGGATTGAGGAGGCACA 2648
Db 19181 GCTTAGTAAACCACTGCACTACTCCCTGACGATTAATCCGTAGTGGATTGAGGAGGCACA 19129

Qy 2649 GAAACCCAGTGCAGTGGAGGGTGTAGTCAAGATCTCAGGATTAATCAATGAGGAGCGGTT 2708
Db 19128 GAAACCCAGTGCAGTGGAGGGTGTAGTCAAGATCTCAGGATTAATCAATGAGGAGCGGTT 19071

Qy 2709 GTCTCTTTATACCCAGCTGTACTAGCCCT---TATACCTGTGTGMYTTTCCAAATACCAGAG 2765
Db 19070 GTTCTCTATACCCAGCTGTACTAGCCCT---TATACCTGTGTGMYTTTCCAAATACCAGAG 19011

Qy 2766 GAAGCAGAGTGGTTTACASTCTGGACCTTMAAGATGCTTCTTCTGTCATCTCTGTGACAT 2825
Db 19010 GAAGCAGAGTGGTTTACASTCTGGACCTTMAAGATGCTTCTTCTGTCATCTCTGTGACAT 18951

Qy 2826 CTTGACTCTCAATTTCTTGTGCTTTTGAAGATACCTTCAAAACCCARCATCTCAACTCACC 2885
Db 18950 CTTGACTCTCAATTTCTTGTGCTTTTGAAGATACCTTCAAAACCCARCATCTCAACTCACC 18891

Qy 2886 TGGACTTTTATACCCCAAGGGTTCAAGGATAGYCCCCATCTATTTGGCCAGGCGATTAGCC 2945
Db 18890 TGGACTTTTATACCCCAAGGGTTCAAGGATAGYCCCCATCTATTTGGCCAGGCGATTAGCC 18831

Qy 2946 CAAGACTTGAGTCATYMTCACTTGGACACTCTTGTCTCTTCRGTAGTGAATGATTTA 3005
Db 18830 CAAGACTTGAGTCATYMTCACTTGGACACTCTTGTCTCTTCRGTAGTGAATGATTTA 18771

Qy 3006 CTTTTRGCVGCCVRYTTTCAGAACCTTGTGCCATCAAGCCACCAAGCRCTCTTMAATTTTC 3065
Db 18770 CTTTTRGCVGCCVRYTTTCAGAACCTTGTGCCATCAAGCCACCAAGCRCTCTTMAATTTTC 18711

Qy 3066 CTCGYACCTGTGGCTACAWGGTTTCCAAACCCARAGCTCARCTCTGCTCAGACGAGGTT 3125
Db 18710 ATGGCCACTGTGGCTACAWGGTTTCCAAACCCARAGCTCARCTCTGCTCAGACGAGGTT 18651

Qy 3126 AATATCTTAGGCTTAARATTTTCAAGGCAACAGGCGCTCAGTGAGGAAATYATCCAG 3185
Db 18650 AATATCTTAGGCTTAARATTTTCAAGGCAACAGGCGCTCAGTGAGGAAATYATCCAG 18591

Qy 3186 CTTATCTGGCTTATCTTCATCTGAAACCTTAAAGCACTAAGGRRTTCTTGGCPTA 3245
Db 18590 CTTATCTGGCTTATCTTCATCTGAAACCTTAAAGCACTAAGGRRTTCTTGGCPTA 18531

Qy 3246 AVAGGYTTCTGCGGAATGATTTCCCGAGGTWTGCGRAAATAGCCAGGYCATTAATATAC 3305
```

[illegible]

Db	17455	GGTGGGACCTCCATTAGTAAATGCCTTATAGAAGGCCCTAGTAGTGGGGTAAATCCCTGT	17396
Qy	4386	GGCAAAACCAAGCCCCCAGTACTCAGCAGGAGAAAACAGAAATGGGGAACCTCACGAGGCAGTT	4445
Db	17395	AGNAACCAGCCCCCAGTACTCAAAGAATAATAGGATGGGAACCTCACGAGGCACATGTT	17336
Qy	4446	TTCCTCCCTCGGAGCGGTAGCCACTGAAGAAGGAAAAAATACTTTTTGCTGCAACTATCC	4505
Db	17335	TCTTCCCTCAGGATGGCTAGCCACCGAAGAAGAAAAAATACTTTTGCTGCAGCTAAC	17276
Qy	4506	AATGGAAATTAACCTTAAACCCCTCATCAACCTCTTCACTTAGGCATCATAGCACCCATC	4565
Db	17275	AATTGAATTAACCTTAAACCCCTTCATCAGACCTTTCACCTTAGGCATCATAGCACCTATC	17216
Qy	4566	ARATGGCCAATCATTTATTTACTTGGACACAGGCCTTTTCAAACTATCAAGCARATAKTCA	4625
Db	17215	AGAGGCCAAATCATTTATTAATCTGGATCAGGACTTTTCAAACATCAAGCAGATAGTCA	17156
Qy	4626	GGGCTGTGAATGTGCCAPARAATAATATCCCCTGCTVATCGCCAAAGCTCTCTTCAGGAR	4685
Db	17155	GGGCTATGAATGTGCCAAAGAAATAATCCCCTGCTTATCGCCAAGCTCTCTCAGGAG	17096
Qy	4686	AACAABAACAGGCCATTACCTGRBARABACTGGCAACT-GATTTTACCACAAGGCCCA	4744
Db	17095	AACAAGAACAGGCCATTACCCAGGAGAGACTGGCACTAGATTTTATCCCAATGCCCCA	17036
Qy	4745	AACCTCAGGATTTCAGTATCTACTAGTCTGGTARATATCTTTACGGTGTGGGCARAGG	4804
Db	17035	AATCTCAGGANTTTCAGTATCTACTAGTCTGGATAGATACITTTCACTGTTGGGAGGAG	16976
Qy	4805	CCTTCCCCTGTAGSACAGAAAAAGGCCCAAGAGGTAAATAAGSCACTAGTTTCATGAANAAT	4864
Db	16975	CCTTCTCTGTAGSACAGAAAAAGGCCCAAGAGCGTAATAAAGGCACTAATTCATGAAATA	16916
Qy	4865	TTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCAGGCCCA	4924
Db	16915	TTCCAGATTCGGACTTCCCTGAGACTTACAGAGTGACAAATGGCCCCGCTTTTCAAGGGCTG	16856
Qy	4925	CAGTAACCCAGGAGGATCCCAAGGCG- - -TTTAGGTATACGATATCACTTACACTGCGCC	4980
Db	16855	CAGTAACCCAGCGATTCGACAGCTTACTTAGGCATACATATCACTTACACTGCGCC	16796
Qy	4981	TGAAGGCCACAGTCTCTCAGGGAAGTTCGAGAAATGAATGAAYAYACTCAAAGGACATCTA	5040
Db	16795	TGGAGGCCACAACTCTCAGGGAAGTTCGAGAAATGAACGAAACACTCAAACAACATCTA	16736
Qy	5041	AAAAAGAACCCAGGAACCCACCTCACATGCCCTGYTCTGTTGGCTTATAGCCTTAAAA	5100
Db	16735	AAAAAGCTAATCAGGAACCCCACTCGATGCCCTGCTCTGTGCCCATAGCCTTATTA	16676
Qy	5101	AGAACTGCAACTTTCCCAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGG	5160
Db	16675	AGAAACAAACCTCCCCCAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGG	16616
Qy	5161	CCCTTCATTAACCAATGACCTTGTGCTTGACCCA-AGACAGGCCAACTTAGTTTGCAGACATC	5219
Db	16615	CCCTTCCTAACCAATGACCTTGTGCTTGACCGAGAGAGGCCCAATTTAGTTTGCAGACATC	16556
Qy	5220	ACTCTCTTAGC AAAATACAAAGTCTTTAAAAANTHCAAGGAACCTATCCCTTGAGAA	5279
Db	16555	ACTCTCTTAGC AAAATATCAACAAGTCTTTAAAAACATTAACAGGGAACCTGTACCCGAGAG	16496
Qy	5280	GAGGGAAGAACTATTTTCCACCWWGTGACATGGTATTAGTCAAGTCCCTTCCTCTAAT	5339
Db	16495	GAGGGAAGAA-----ATTCACACCCTGGTAACATGGTATTAGTCAAGTCCCTTCCCTCTAAG	16439
Qy	5340	TCCCCTCCCTAGATACATCTGGGAAGGACCCCTACCCAGTCAATTTTATYTATCCCAACT	5399
Db	16438	TCCCACCCCTAGATACATCTGGGAAGGACCCCTACCCAGTCAATTTTATCTACCTCTAGT	16379
Qy	5400	CGGTTTAAAGTGCGTGGAGTGCTTGGATATCATCACATTTAGTCAAACTCTTGATA	5459
Db	16378	GCAGTTAAAGTGTCTGGAGTGAGTCTTAGACACA-----TCAAACTCTTGATA	16330

```
QY 5460 CTGCCAAGAACCTGAAATCCAGGAGCAACCGTAGCTATTCTCTGTGAACCTCTAGAG 5519
D 16329 CCGCAAGAGAACCAAAATTCAGGAGCAACCGTAGCTATTCTCTGTGAACCTCTAGAG 16270
QY 5520 GATTGGCGCTGCTCTTCAAAACAACACGAGAGAAAGTAATCAAAATCAATAATCCCC 5579
D 16269 GATCTGCACCTGCTCTTCAAGCGACAACCTGCGAGGAAGTAATCGTAGAGCTCC 16210
QY 5580 CATGSCCTCCCTTATCATATTTCTCTKTASTGTTSTTYA-CCCTSTTTCACTCA 5638
D 16209 --ATGGGCCCCCGCTCATATTTCTTTACTTGTGTCTTACCCCTTTCACTATCA 16152
QY 5639 CTGCACCCCTCCATGCGCTGTATGACCACTAGCTCCCTYACMAGAGTTTCTATGGA 5698
D 16151 CCTCACCCCTCCATGCGCTGTATGACCACTAGCTCCCTTACCAAGAGCTTCTATGGA 16092
QY 5699 GAATGACGCGTCCCGGAAATATGATGCCCATCGTATAGGAGTCTTTSTAAGGGAACCC 5758
D 16091 GAATGCGCTTCCAGAAATATGATGCCCATGTATAGGAGTTTCTTAAGGGAACCC 16032
QY 5759 CCACCTTCACTGCCCAACCACTATGCGCCGCAACTGC 5796
D 16031 CCACCTTCACTGCCCAACCACTATGCGCCGCAACTTC 15994
```

## RESULT 7

```
US-09-175-928-3
; Sequence 3, Application US/09175928A
; Patent No. 6312921
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Metberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Mi, Sha
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6006B.AJ172A
; CURRENT APPLICATION NUMBER: US/09/175,928A
; CURRENT FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2946
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-175-928-3
```

```
Query Match 38.6%; Score 2888.6; DB 3; Length 2946;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 2872; Conservative 49; Mismatches 3; Indels 1; Gaps 1;

QY 4658 CTGCTVATCGCAAGCTCCTTCAGGARAACAARAACAGGCCATTACCCCTGPARARAC 4717
D 6 CTGCTTATCGCAAGCTCCTTCAGGAGAACAAGAACAGGCCATTACCCCTGGAGAGAC 65
QY 4718 TGGCACTGATTTTACCACAAGCCCAACCTCAGGGATTTTCAGTATCTACTAGTCTGGG 4777
D 66 TGGCACTGATTTTACCACAAGCCCAACCTCAGGGATTTTCAGTATCTACTAGTCTGGG 125
QY 4778 TATATACCTTTACGGGTTGGGAGAGGCTTCCCTGTAGGACAGAAAGGCCCAAGNG 4837
D 126 TATATACCTTTACGGGTTGGGAGAGGCTTCCCTGTAGGACAGAAAGGCCCAAGAGG 185
QY 4838 TAATAAGGCACTAGTTTCATGAATAATTCACAGATTTCGAGCTTCCCGAGGCTTACAGA 4897
D 186 TAATAAGGCACTAGTTTCATGAATAATTCACAGATTTCGAGCTTCCCGAGGCTTACAGA 245
QY 4898 GTGACAATAGCCCTGCTTCCAGGCCACAGTAACCCAGGGAGTATCCAGGGCGTTAGGTA 4957
```

```
Db 246 GTGACAATAGCCCTGCTTCCAGGCCACAGTAACCCAGGGAGTATCCAGGCGTTAGGTA 305
QY 4958 TACGATATCACTTACACTGCGCTTGAAGGCCACAGTCTCAGGGAAGTTCAGAAATGA 5017
D 306 TACGATATCACTTACACTGCGCTTGAAGGCCACAGTCTCAGGGAAGTTCAGAAATGA 365
QY 5018 ATGAAAVACTCAAAAGGACATCTAAAAAGCAAAACCCAGGAAACCCACTCACAATGGCGCTG 5077
D 366 ATGAAACACTCAAAAGGACATCTAAAAAGCAAAACCCAGGAAACCCACTCACAATGGCGCTG 425
QY 5078 YTCGTGTCCTATAGCCCTTAAAGATCTGCACTTTCCCAAAAAGCAGGACTTAGCC 5137
D 426 CTCGTGTGCTATAGCCCTTAAAGATCTGCACTTTCCCAAAAAGCAGGACTTAGCC 485
QY 5138 CATAACGAAATGCTGTATGGAAGGCCCTTCAATAACCAATGACCTTGTGCTTGACCCCAAGAC 5197
D 486 CATAACGAAATGCTGTATGGAAGGCCCTTCAATAACCAATGACCTTGTGCTTGACCCCAAGAC 545
QY 5198 AGCCAACTTATGTTGACAGATCACTCTCTTAGCAAAATATCAACAAGTTCTTAAACAATT 5257
D 546 AGCCAACTTATGTTGACAGATCACTCTCTTAGCAAAATATCAACAAGTTCTTAAACAATT 605
QY 5258 ACAAGGAACCTATCCCTGAGGAAGGGAAGAACTATTCCACCWWGTGACATGGTATT 5317
D 606 ACAAGGAACCTATCCCTGAGGAAGGGAAGAACTATTCCACCWWGTGACATGGTATT 665
QY 5318 AGTCAAGTCCCTTCYCTCTAATTTCCCATCCCTAGATACATCTCTGGGAAGACCCCTACCC 5377
D 666 AGTCAAGTCCCTTCYCTCTAATTTCCCATCCCTAGATACATCTCTGGGAAGACCCCTACCC 725
QY 5378 AGTCAATTTATYACCCCCAACTCGCGTTAAAGTGGTGGAGTGGAGTCTTTGGATACATCA 5437
D 726 AGTCAATTTATYACCCCCAACTCGCGTTAAAGTGGTGGAGTGGAGTCTTTGGATACATCA 785
QY 5438 CACTTGAGTCAAAATCCTGGATCTGCTCAAGGAACCTGAAATTCAGGAGACAACGCTAG 5497
D 786 CACTTGAGTCAAAATCCTGGATCTGCTCAAGGAACCTGAAATTCAGGAGACAACGCTAG 845
QY 5498 CTATTCCTGTGAACCTCTAGAGGATTTGGCCCTGCTCTTCAAAACAACACAGGAGAAA 5557
D 846 CTATTCCTGTGAACCTCTAGAGGATTTGGCCCTGCTCTTCAAAACAACACAGGAGAAA 905
QY 5558 GTAACTAAAATCATAAATCCCCATGSSCTCCCTTATCATATTTTTCTCTKTASTGTT 5617
D 906 GTAACTAAAATCATAAATCCCCATGSSCTCCCTTATCATATTTTTCTCTKTASTGTT 964
QY 5618 TTTTACCCCTCTTCACTCTCACTGCAACCCCTCCATGCGCTGTATGACCAAGTAGCTCCC 5677
D 965 TTTTACCCCTCTTCACTCTCACTGCAACCCCTCCATGCGCTGTATGACCAAGTAGCTCCC 1024
QY 5678 CTYACCMAGAGTTTCTATGGAGAAATCGAGGTCCTCCGGAATATTTGATGCCCTTCGTATA 5737
D 1025 CTYACCMAGAGTTTCTATGGAGAAATCGAGGTCCTCCGGAATATTTGATGCCCTTCGTATA 1084
QY 5738 GGAGTCTTTSTAAGGAAACCCCACTTCACTGCCCAACCCATATGTCGCCGCAACTGCT 5797
D 1085 GGAGTCTTTSTAAGGAAACCCCACTTCACTGCCCAACCCATATGTCGCCGCAACTGCT 1144
QY 5798 ATCACTCTGCACTCTTTGCACTGCAAAATACTCATTTATGGAACAGGAAAAATGATTA 5857
D 1145 ATCACTCTGCACTCTTTGCACTGCAAAATACTCATTTATGGAACAGGAAAAATGATTA 1204
QY 5858 ATCTCTAGTTGCTCTGAGGACTTGGAGTCACTGCTGTGTTGAGTCTTCACTCCCAACTG 5917
D 1205 ATCTCTAGTTGCTCTGAGGACTTGGAGTCACTGCTGTGTTGAGTCTTCACTCCCAACTG 1264
QY 5918 GTATGTCGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAATGATTAAGAGATTA 5977
D 1265 GTATGTCGATGGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAAATGATTAAGAGATTA 1324
QY 5978 TCTCCCAACTCAACCGGGGTACATGGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAA 6037
```



Db 1325 TCTCCCACTACACCGGGTACATGGCACTCTAGCCCTTACAAAGGACTAGATCTCTCA 1384  
 Qy 6038 AACTACATGAAGAACCTTCGCTACCCATACCTGCTGGTAAGCCTATTTTAATACACCTCA 6097  
 Db 1385 AACTACATGAAGAACCTTCGCTACCCATACCTGCTGGTAAGCCTATTTTAATACACCTCA 1444  
 Qy 6098 CTGGGCTCCATGAGTCTCGGCCCAAAACCTTACTAACTGTTGGATATGCTCCCTCGA 6157  
 Db 1445 CTGGGCTCCATGAGTCTCGGCCCAAAACCTTACTAACTGTTGGATATGCTCCCTCGA 1504  
 Qy 6158 ACTTCARGGCATATGTTTCAATCCCTGTACCTGAACAAATGGAACTTTCAGCACAGAAA 6217  
 Db 1505 ACTTCARGGCATATGTTTCAATCCCTGTACCTGAACAAATGGAACTTTCAGCACAGAAA 1564  
 Qy 6218 TAAACACCACTTCCTGTTTGTAGGACCTCTGTTTCCAAATGTGGAATTAACCCATACCT 6277  
 Db 1565 TAAACACCACTTCCTGTTTGTAGGACCTCTGTTTCCAAATGTGGAATTAACCCATACCT 1624  
 Qy 6278 CAAACCTCAGCTGTGTAATTTAGCAATACTACATACACCAACCACTCCCAATGCATCA 6337  
 Db 1625 CAAACCTCAGCTGTGTAATTTAGCAATACTACATACACCAACCACTCCCAATGCATCA 1684  
 Qy 6338 GTGGGTAACTCTCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTG 6397  
 Db 1685 GTGGGTAACTCTCTCCACACAAATAGTCTGCTACCTCAGGAATATTTTGTCTGTG 1744  
 Qy 6398 GTACCTCAGCTATCGTGTGTTGAAATGGCTCTTCAGAACTATGTCCTCTCATCT 6457  
 Db 1745 GTACCTCAGCTATCGTGTGTTGAAATGGCTCTTCAGAACTATGTCCTCTCATCT 1804  
 Qy 6458 TAGTGGCCCCATGACCTACCTGACCAAGATTTATACAGTTATGTCATATCTAGC 6517  
 Db 1805 TAGTGGCCCCATGACCTACCTGACCAAGATTTATACAGTTATGTCATATCTAGC 1864  
 Qy 6518 CCGGCAACAAAGAGTACCCATTCTCTCTTTGTTATAGGAGCAGGAGTCTAGTGCAC 6577  
 Db 1865 CCGGCAACAAAGAGTACCCATTCTCTCTTTGTTATAGGAGCAGGAGTCTAGTGCAC 1924  
 Qy 6578 TAGGTACTGGCATGTCGCGTATCACAACTCTACTCAGTTCTACTACAACTATCTCAAG 6637  
 Db 1925 TAGGTACTGGCATGTCGCGTATCACAACTCTACTCAGTTCTACTACAACTATCTCAAG 1984  
 Qy 6638 AACTAAATGGGACATGGAACGGTCCGACCTCCCTGCTCAGCTTGCAGAGTCACTTA 6697  
 Db 1985 AACTAAATGGGACATGGAACGGTCCGACCTCCCTGCTCAGCTTGCAGAGTCACTTA 2044  
 Qy 6698 ACTCCCTAGCAGCAGTGTCTTTCRAAATCGAAGACTTTAGACTGCTTAAACCGCTGARA 6757  
 Db 2045 ACTCCCTAGCAGCAGTGTCTTTCRAAATCGAAGACTTTAGACTGCTTAAACCGCTGARA 2104  
 Qy 6758 GAGGGGAACTGTTTATTTTAGGGGAAGAAATGCTGTTTATTGTTAATCAATCCGAA 6817  
 Db 2105 GAGGGGAACTGTTTATTTTAGGGGAAGAAATGCTGTTTATTGTTAATCAATCCGAA 2164  
 Qy 6818 TGTCTACTGAGAAGTTAAGAAATTCAGATCGAATACAACTAGAGCAGAGAGCTTC 6877  
 Db 2165 TGTCTACTGAGAAGTTAAGAAATTCAGATCGAATACAACTAGAGCAGAGAGCTTC 2224  
 Qy 6878 GAAACACTGGACCTCGGCGCTCTCAGCCATGAGTCCCTGATTTCTCCCTCTTAG 6937  
 Db 2225 GAAACACTGGACCTCGGCGCTCTCAGCCATGAGTCCCTGATTTCTCCCTCTTAG 2284  
 Qy 6938 GACCTCTAGCAGCTAATAATGCTACTCTCTTTTGGACCTGTATCTTTTACCTCTTG 6997  
 Db 2285 GACCTCTAGCAGCTAATAATGCTACTCTCTTTTGGACCTGTATCTTTTAACTCTTG 2344  
 Qy 6998 TTAACCTTGTCTCTCAGATCGAAGCTGTAACTACAAATGGAGCCCAAGATGACGT 7057  
 Db 2345 TTAACCTTGTCTCTCAGATCGAAGCTGTAACTACAAATGGAGCCCAAGATGACGT 2404  
 Qy 7058 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCTGTAGCCCAAGATCTGATGTTA 7117  
 Db 2405 CCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCTGTAGCCCAAGATCTGATGTTA 2464

Qy 7118 ATGACATCAAGGACACCCCTCTGAGGAAATCTCAGCTGCACAACTCTTACTACGCCCA 7177  
 Db 2465 ATGACATCAAGGACACCCCTCTGAGGAAATCTCAGCTGCACAACTCTTACTACGCCCA 2524  
 Qy 7178 ATTACAGAGGAAGCAGTGTAGAGCGGCTGTCGGCCAACTCCCAACAGCAGCTTAGGTTT 7237  
 Db 2525 ATTACAGAGGAAGCAGTGTAGAGCGGCTGTCGGCCAACTCCCAACAGCAGCTTAGGTTT 2584  
 Qy 7238 CTTGTTGAGATGGGGGACTGAGAGCAGGACTAGCTGATTTCTTAGGCTGATTAAGAT 7297  
 Db 2585 CTTGTTGAGATGGGGGACTGAGAGCAGGACTAGCTGATTTCTTAGGCTGATTAAGAT 2644  
 Qy 7298 CCYTAAGCCTAGTGGGAAGTGAACCATCCACTTTAAACACCGGCTTCGCACTTAG 7357  
 Db 2645 CCCTAAGCCTAGTGGGAAGTGAACCATCCACTTTAAACACCGGCTTCGCACTTAG 2704  
 Qy 7358 YTCACACCTGACCAATCAGAGAGCTCACTAAATGCTAAATAGGCAAGACAGGAGTTAA 7417  
 Db 2705 CTCACACCTGACCAATCAGAGAGCTCACTAAATGCTAAATAGGCAAGACAGGAGTTAA 2764  
 Qy 7418 AGAATATGCCAATCATYTTATTGCTGAGAGCAGCAGGAGGACAATGATCGGATATA 7477  
 Db 2765 AGAATATGCCAATCATCTATTGCTGAGAGCAGCAGGAGGACAATGATCGGATATA 2824  
 Qy 7478 AACCCAACTGTCGAGCGGCAACCGCAACCCCTTTGGTCCCTCTTGTATGGCA 7537  
 Db 2825 AACCCAACTGTCGAGCGGCAACCGCAACCCCTTTGGTCCCTCTTGTATGGCA 2884  
 Qy 7538 GCTCTGTTTTCATGCTATTTTCACTCTATTAATCTTGCARCTGCR 7582  
 Db 2885 GCTCTGTTTTCATGCTATTTTCACTCTATTAATCTTGCARCTGCA 2929

RESULT 8

US-09-949-016-13002/c  
 ; Sequence 13002, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 13002  
 ; LENGTH: 168394  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; FEATURE:  
 ; NAME/KEY: misc.feature  
 ; LOCATION: (1)-(168394)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-949-016-13002

Query Match 34.9%; Score 2615.4; DB 3; Length 168394;  
 Best Local Similarity 85.1%; Pred. No. 0;  
 Matches 3133; Conservative 134; Mismatches 355; Indels 58; Gaps 33;  
 Qy 1012 TTCCCCCAATTATGACCCCTCCCAAGCAGTGGGAGGAGAAATTCGGCCAGCCAGATG 1071  
 Db 29005 TCCCCCAAGGCAAAACACCCCTTAAATGTATCTGGAGAAATTCGGCTTGTACAGTG 28946  
 Qy 1072 CATGTGCTTTTCTTCTCCAGACTTAAAGCAAAATAAACAGACTTAGGTAAATTTCTCA 1131



Db	28945	TATGTACCTTTTCCCTGTGACAGCTTGAAGCAAAATTAATAATAGACCTAGGTAAATTTCTCA	28886
Qy	1132	GATAAYCCTGATGCTATATGRTGTTTTTACAAGGGTTAGGACAAATTTTGTGATCTGACA	1191
Db	28885	GATAAACCTGTGTGCTATATGATGCTTTTACAAGGGTTAGGACAAATCTTTGATCTGACA	28826
Qy	1192	TGGAGAGATATATGTCACTGCTAAATACAGACATAACCCCAATCAGAGAGAGTGCAC	1251
Db	28825	TGGAGAGATATA-ATGTTACTGCTAGATCAGACACTAATCCCAATCAGAGAGAGTGTGC	28767
Qy	1252	CATAACTGCGAGCTGAGRGTTTGGCGATCTCTGATATCTCAGTCAGTCGAATGGATANGG	1311
Db	28766	CATAACTGCGAGCTGAGAGTTTGGTGATCTCTGATATCTCAGTCAGTCGAATG--ATAGG	28709
Qy	1312	ATGACAAAGAGAGAAAGANAATGATTTCCCAAGGCGAGCAGTCAGTCTCCAGCTCTASA	1371
Db	28708	ATGACAAAG-AGGAAAGAGCAAAATTTCCCAAGGCGAGCAGTCTCCAGCTCTASA	28650
Qy	1372	CCCTCATTTGGGACACAGAAATCAGTAAACATGGAGAGTTGGTCTGCGAGACATTTGCTAA	1431
Db	28649	CCCTCATTTGGG-----ACACAGAAATCAGACATGGAATTTGGTCCACAGACATTTGCTAA	28595
Qy	1432	CTTGTGTGCTASAGGACTAAGGAAACTAGGAAARTCTAYGAAATTAATCTCAATGATG	1491
Db	28594	CTTGTGTGCTA-AAGGACTAAGGAAACTTAGAAG-AGGCTATAGATTAATTCAGTGATG	28537
Qy	1492	TCCACCATACACAGGGGAGGAGAAATCTCTAATCTGCTTTCTGGAGAGACTAAGGGA	1551
Db	28536	TTCACTATATACACA-AGGAAAGGAGAAATCTCTGCTGCTTTCTGTAGAGACTAAGGGA	28478
Qy	1552	GGCATTTGAGGAAGCTGCTCTCTGTCACCTGACTCTCTGGAAGGCCCACTTAATCTTAA	1611
Db	28477	GGCATTTGAGGAAGC-AAACTCTCTGTCACCTGACTCTGTTGAAGGCCCACTTAATCTTAA	28419
Qy	1612	GGCTAAGTTTATCACTCAGTCAGTCGACATTAG-AAAAAACTTCAAAAGTCTGCCGT	1670
Db	28418	GGTTAAGTTTATCACTCAGTCAGTCGACATTAGAAAAAACTTTAAAGTCTCACTT	28359
Qy	1671	AGSCCGGAGCAAACTTAGAAACCTTATGAACTTGGCAACTGCGTTTATTAATAG	1730
Db	28358	AGSCCGGAGCAAACTTAGAAACCTTATGAACTTGGCAACTGCGTTTATTAATAG	28299
Qy	1731	AGATCAGGAGGAGCGCGGAAACAGGACAAACGGGATTAATAAAAGGCCACCGCTTAG	1790
Db	28298	AGATCAGGAGGAGCGCGGAGAA-CAGGATAACGGGA-TAAAGAAAGGCCACCACTTAG	28240
Qy	1791	TCATGACCCCTCAGGCAAG-TGGACTTTTGGAGGCTCTGGAAGGGAAGGCTGGGCAAT	1849
Db	28239	TCATGACCCCTCAGGCAAGCAAGCTTGGAGGCTCTGGAACAGGGAAGTCTGGACAAAT	28180
Qy	1850	TGAATGCCCTAATAGGCTTGTCTCCAGTGGGCTTCAAGGACACTTTAAAAAGATTGT	1909
Db	28179	TGAATGCCCTAATAGGCTTGTCTCCAGTGGGCTTCAAGGACACTTTAAAAAGATTGT	28120
Qy	1910	CAAAGTAGAAGTAAGCC-GCCCTTGTGTCATGCCCTTATTTCAAGGGAATCACTGGAA	1968
Db	28119	TCATATAGAAATTAAGCAACCCCTCTGTCATGGCCCTTATGTCAAGGGAATCACTGGAA	28060
Qy	1969	GGCCCACTGCCCGAGGGGCAAAAGGCTTTTGTAGTCAGAGGCCACTTAAACAGATGATCCA	2028
Db	28059	GGCCCACTGCCCGAGGGGAGAGGCTCTCTGAGTCAGAGGCCACTTAAACAGATGATCCA	28000
Qy	2029	GCAGCAGGACTGAGGCTGCTGGGGCAAGCGGCATCCCATGCCATCACCCCTCACAGAGCC	2088
Db	27999	GCAGCAGGACTGAGGCTGCTGGGGCAAGCGGCATCCCATGCCATCACCCCTCACAGAGCC	27940
Qy	2089	CTGGGTATGCTTGA-CAATTTAGGGGCGAGGAAGT---TGTCTCTGGACA-CTGGTGCCTG	2145
Db	27939	CCGGGTATGCTTAA-CAATTTAGGGGCGAGGAGTTAACTGTCTCTGGACA-CTGGGCTGCTG	27880
Qy	2146	CTTCTTAGCTTACTCTTCTGCTGGGCAAACTGTGCTCTCCAGATCTGCTCACTATTCTGAG	2205
Db	27879	CTTCTAGCTTACTCTCTCTGCTGGGCAAACTGT-CTCCAGATCTGCTCACTATCCAAGG	27821
Qy	2206	GGGTCCTNTAAGACGGGACGTCACTAGATACCTTTTTCGCCACCACTAAGTTATGAACCTGG	2265
Db	27820	GG-----TCCTGGGACAGCCACTAGATAC-TTCTCCAGGCACTAAGCTGTG-ACTGG	27770
Qy	2266	GGAGCTTTATTTTTCACATGCTTTTCTAATATATGCTTGAAGGCCCACTACCTTGTGTA	2325
Db	27769	GGAACTTTATTTTTCACATGCTTTTCTAATATATGCTTGAAGGCCCACTACCTTGTGTA	27710
Qy	2326	GGGAGAGACATTTCTAGCAAAAGCAGGGGCGCATTTATACACCTGAACTAGAGAGAGGAACA	2385
Db	27709	GGGAAAGACATTTCTAACAAGCAGGGACCATTTATACACCTGAACTAGAGAGAGGAACA	27650
Qy	2386	CCGTTTGTGTGTCCTGCTTGGAGGAGGAAATTAATCTCTGAAAGTCTGGGCAACAGAGG	2445
Db	27649	CCGTTTGTGTGT-CCCTGCTTGAAGAGGAAATTAATCTCTAAGTCTGGGCAACAGAGG	27591
Qy	2446	ACAATATGGACGAGCAAAAGAAATGCCGCTCTCTCTAGTTAAACTAAGAGATTTCCACCT	2505
Db	27590	ACAATATGGACGAG-CAAAATTAATGCCGCTCTCTCTAGTTAAACTAAGAGATTTCCACCT	27532
Qy	2506	CTTTTCCCTACCAAGGACGTACCCCTCAGACCCCAAGGCCCAACAGGATTTCCAAAGA	2565
Db	27531	CTTTTCCCTACCAAGGACGTACCCCTTAGACCCGAGGCCCAACAGGACTTCCAAAAA	27472
Qy	2566	TTGTTAAGGACTTAAAGGCCCAAGGCTTAGTAAAAACCATGATTAATCTCCCTGCAGTAAT	2625
Db	27471	TTGTTAAGGACCTTAAAGGCCCTACTCTAGTAAAAACCATGATTAATCTCCCTGCAGTAAT	27412
Qy	2626	CCGTAGTGGATTGAGGAGGACAGAAACCCAGTGGGACAGTGGAGGGTTAGTCAAGATCT	2685
Db	27411	C-----AATTTTAGGAGTAAAGAGCCCAATAGACAGTGA-GGTTAGTCAAGATGT	27360
Qy	2686	CAGGATTATCAATGGAGGCGTGTCTTTTATACCCAGCTGTACCTAGCCCTTATACITG	2745
Db	27359	CAGGATTATCAGT-GAGGCTGTCTTCTCTATACCCAGCTGTACCTTACCTTCTACTC	27301
Qy	2746	TGMYTTCCCAATACACAGAGGAGCAGAGTGGTTTACASTCTTGGACCTTMMAGATGCCT	2805
Db	27300	TGCTTTCCCAATACACAGAGGAGCAGAGTGGTTTACAGTCTCTGGACCTTAAAGATGCCT	27241
Qy	2806	TCCTTGTGATCCCTGTATCATCTGACTCTCAATCTTGTGTGCTTTTGAAGATCTTCAA	2865
Db	27240	TTTCTGTGATCCCTGTATCATCTGACTCTCAATCTTATTTGCTTTTAAAGATGTCAA	27181
Qy	2866	ACCCAFCACTCAACTCAGCTGCACTRTTTTATCCCAAGGTTTCAAGGATAGTCCCATC	2925
Db	27180	ACCCATGTCTCAACTCAGCTGCACTRTTTTATCCCAAGGTTTCAAGGATAGTCCCATC	27121
Qy	2926	TATTTGGCCAGGCAATTTAGCCCAAGACTTGAGYCATYMTCACTGGAACACTCTTTGTCC	2985
Db	27120	TATTTGGCCAGGCAATTTAGCTCAAGACTTGAGACAGTTCATACCTGGAACACTCTTTGTCC	27061
Qy	2986	TTTCGTAGTGTGATGATTTACTTTTGTGCGCCYRTT-CAGAAAACTTGTGCCATCAAGCC	3044
Db	27060	TTTGGTACATGATAATCTACTTCTAGCTGCTGTTTGAAGAAAACTTGTGCCATCAAGCC	27001
Qy	3045	ACCCAGGCTCTTMAATTTCTCGYACCTGTGGCTACAWGGTTTCCAAAASARAGCT	3104
Db	27000	ACCCAGGCTCTTMAATTTCTCGYACCTGTGGCTACAWGGTTTCCAAAASARAGCT	26941
Qy	3105	CARCTCTGTCTACAGCAGGTTAAATTAATTTAGGCTTAATTAATCCAAAGGCCACCGGCC	3164
Db	26940	CAGCTCTGTCTACAGCAGGTTAAATTAATTTAGGCTTAATTAATCCAAAGGCCACCGGCC	26881
Qy	3165	CTCAGTAGGAAATGATCCAGCTTATCTGGCTTATCTCTCATCYCAAAACCTTAAAGCAA	3224
Db	26880	CTCAGTAGGAAATGATCCAGCTTATCTGGCTTATCTCTCATCCCAAAACCTTAAAGCAA	26821
Qy	3225	CTAAGRGRTTCTTGGGRTTAAGGTTCTGCCGAATGATTTCCCAAGGTTGGGCA	3284
Db	26820	ATAAAG-GTTCTTGGGATAACAGGCTTCTGCCAAATATGATTT-CCCAAGTACGGCAA	26763



Db 304 GTATACGATATCACTTACCTGCGCTGAAGGCCACAGTCTCAGGGNAGGTGAGAAA 363  
Qy 5015 TGAATGAAYACTCAAGAGCATCTAAAAAGGAAACCCAGGAAACCCACCTCACTGGC 5074  
Db 364 TGAATGAACACTCAAGAGCATCTAAAAAGGAAACCCAGGAAACCCACCTCACTGGC 423  
Qy 5075 CTGTCCTGTTGCGCTATAGCCCTTAAAAAGAAATCTGCAACTTTCCCAAAAAAGCAGGACTTA 5134  
Db 424 CTGCTCTGTTGCGCTATAGCCCTTAAAAAGAAATCTGCAACTTTCCCAAAAAAGCAGGACTTA 483  
Qy 5135 GCCCATACGAAATGCTGTATGGAAGGCCCTTCAATACCAATGACCTTTGTGCTTGACCCAA 5194  
Db 484 GCCCATACGAAATGCTGTATGGAAGGCCCTTCAATACCAATGACCTTTGTGCTTGACCCAA 543  
Qy 5195 GACAGCAACTTAGTTCAGACATCACTCTCTAGCCAAATATCAACAAGTTCTTTAAAC 5254  
Db 544 GACAGCAACTTAGTTCAGACATCACTCTCTAGCCAAATATCAACAAGTTCTTTAAAC 603  
Qy 5255 ATTACAAGGAACCTATCCCTGAGAAGAGGAAAGAACTATTCCACCCTGAGCATGGT 5314  
Db 604 ATTACAAGGAACCTATCCCTGAGAAGAGGAAAGAACTATTCCACCCTGAGCATGGT 663  
Qy 5315 ATTAGTCAAGTCCCTTCYCTTAATCCCATCCCTAGATACATCTCTGGGAAGGACCCCTA 5374  
Db 664 ATTAGTCAAGTCCCTTCYCTTAATCCCATCCCTAGATACATCTCTGGGAAGGACCCCTA 723  
Qy 5375 CCAGTCATTTATYATACCCCAACTGCGGTAAAGTGGCTGGAGTGAGTCTTTGGATACA 5434  
Db 724 CCAGTCATTTATYATACCCCAACTGCGGTAAAGTGGCTGGAGTGAGTCTTTGGATACA 783  
Qy 5435 TCACACTTGAATCAAAATCCCTGAGATCTGCGCTGCTCTTCAAAACAAACACGAGG 5494  
Db 784 TCACACTTGAATCAAAATCCCTGAGATCTGCGCTGCTCTTCAAAACAAACACGAGG 843  
Qy 5495 TAGCTATTCTGTGAACTCTAGAGGATTTGCGCTGCTCTTCAAAACAAACACGAGG 5554  
Db 844 TAGCTATTCTGTGAACTCTAGAGGATTTGCGCTGCTCTTCAAAACAAACACGAGG 903  
Qy 5555 AAGTAACTAAATCAATAATCCCATGAGGCTTCCCTTATCATATTTTCTCTKTAAGT 5614  
Db 904 AAGTAACTAAATCAATAATCCCATGAGGCTTCCCTTATCATATTTTCTCTKTAAGT 962  
Qy 5615 TTSTTTTACCTSTTTTACCTCTCACTGCAACCCCTCCATGCGCTGTATGACAGTAGCT 5674  
Db 963 TTCTTTTACCTCTTTTCACTCTCACTGCAACCCCTCCATGCGCTGTATGACAGTAGCT 1022  
Qy 5675 CCCTTACCMAGAGTTTCTATGAGAAATGAGAGTCCCGGAAATATTGATGCCCCATCGT 5734  
Db 1023 CCCTTACCMAGAGTTTCTATGAGAAATGAGAGTCCCGGAAATATTGATGCCCCATCGT 1082  
Qy 5735 ATAGGAGTCTTTTAAAGGAAACCCCACTTCACTGCCCCACACCCATATGCCCGCAACT 5794  
Db 1083 ATAGGAGTCTTTTAAAGGAAACCCCACTTCACTGCCCCACACCCATATGCCCGCAACT 1142  
Qy 5795 GCTATCACTCTGCCACTCTTTGATGATGCAATCAATATGAGAGGAAATGA 5854  
Db 1143 GCTATCACTCTGCCACTCTTTGATGATGCAATCAATATGAGAGGAAATGA 1202  
Qy 5855 TTAATCTAGTCTGCTGGAGGACTTGGAGTCACTGCTGTTGGAATCTTCACTTCACTTCA 5914  
Db 1203 TTAATCTAGTCTGCTGGAGGACTTGGAGTCACTGCTGTTGGAATCTTCACTTCACTTCA 1262  
Qy 5915 CTGGTATGCTGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAAAAGAG 5974  
Db 1263 CTGGTATGCTGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAACATGTAAAAAGAG 1322  
Qy 5975 TAATCTCCCAACTCAACCGGGTACATGGCACTCTAGCCCTTACAAAGGACTAGATCTCT 6034  
Db 1323 TAATCTCCCAACTCAACCGGGTACATGGCACTCTAGCCCTTACAAAGGACTAGATCTCT 1382  
Qy 6035 CAAAACTACATGAACCCCTCCGTACCCATACCTGCTGGTGAAGCCTTATTAAATACCACCC 6094  
Db 1383 CAAAACTACATGAACCCCTCCGTACCCATACCTGCTGGTGAAGCCTTATTAAATACCACCC 1442

Qy 6095 TCACTGGGCTCCATGAGGTCTCGGCCCAAAAAACCTACTAACTGTGTGATATGCTCCCCC 6154  
Db 1443 TCACTGGGCTCCATGAGGTCTCGGCCCAAAAAACCTACTAACTGTGTGATATGCTCCCCC 1502  
Qy 6155 TGAACCTTCAAGCCCATATGTTTCAATCCCTGCTGCTGAACTGAAACAACTTTCAGACAG 6214  
Db 1503 TGAACCTTCAAGCCCATATGTTTCAATCCCTGCTGCTGAACTGAAACAACTTTCAGACAG 1562  
Qy 6215 AATAAACAACCACTTCCGTTTATAGTAGGACCTCTTGTTCCTCAATSTCGAAATTAACCCATA 6274  
Db 1563 AATAAACAACCACTTCCGTTTATAGTAGGACCTCTTGTTCCTCAATCTGAAATTAACCCATA 1622  
Qy 6275 CCTCAAAACCTCACTGCTGTAATAATTTAGCAATATACATACAACAACCTTCCCAATGCA 6334  
Db 1623 CCTCAAAACCTCACTGCTGTAATAATTTAGCAATATACATACAACAACCTTCCCAATGCA 1682  
Qy 6335 TCAGGTGGGTAACTCTCCACACAATAAGTCTGCTCCTACCTCAGGAAATATTTTGTCT 6394  
Db 1683 TCAGGTGGGTAACTCTCTCCACACAATAAGTCTGCTCCTCAGGAAATATTTTGTCT 1742  
Qy 6395 GTGCTACCTCAGGCTATCGTTTGAATGGCTCTTCAAGATCTATGCTGCTCTCTCAT 6454  
Db 1743 GTGCTACCTCAGGCTATCGTTTGAATGGCTCTTCAAGATCTATGCTGCTCTCTCAT 1802  
Qy 6455 TCTTAGTGCCCCYATGRCATCTACACTGAAACAAGATTTATACAGTTATGTCTATCTA 6514  
Db 1803 TCTTAGTGCCCCYATGRCATCTACACTGAAACAAGATTTATACAGTTATGTCTATCTA 1862  
Qy 6515 AGCCCCGCAACAAAAGAGTACCCATTTCTTCTTTGTATAGGAGCGAGTGTAGGTG 6574  
Db 1863 AGCCCCGCAACAAAAGAGTACCCATTTCTTCTTTGTATAGGAGCGAGTGTAGGTG 1922  
Qy 6575 CACTAGTACTGCTGATTTGGGGTATCAACCTCTACTCAGTTCTACTCAAACTATCTC 6634  
Db 1923 CACTAGTACTGCTGATTTGGGGTATCAACCTCTACTCAGTTCTACTCAAACTATCTC 1982  
Qy 6635 AAGAACTAAATGGGACATGGAACGGGTGCGGACTCCCTGGTCACTTTCGAAGATCAAC 6694  
Db 1983 AAGAACTAAATGGGACATGGAACGGGTGCGGACTCCCTGGTCACTTTCGAAGATCAAC 2042  
Qy 6695 TTAACCTCCCTAGCAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGACTGCTTAA 6754  
Db 2043 TTAACCTCCCTAGCAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGACTGCTTAA 2102  
Qy 6755 ARAGAGGGGAACTCTGTTTATTTTAGGGGAAAGTCTGTTATTTAGTTAACTCAATCCG 6814  
Db 2103 ARAGAGGGGAACTCTGTTTATTTTAGGGGAAAGTCTGTTATTTAGTTAACTCAATCCG 2162  
Qy 6815 GAATCTGCTCACTGAGAAAGTTTAAAGAAATTCAGATCGAATACAACTGAGCAGAGAGC 6874  
Db 2163 GAATCTGCTCACTGAGAAAGTTTAAAGAAATTCAGATCGAATACAACTGAGCAGAGAGC 2222  
Qy 6875 TTGCAAAACCTGGAACCTCGGGGCTCTCAGCCATGAGTGGCTGGATTTCTCCCTTCT 6934  
Db 2223 TTGCAAAACCTGGAACCTCGGGGCTCTCAGCCATGAGTGGCTGGATTTCTCCCTTCT 2282  
Qy 6935 TAGGACCTTACGAGCTAATATTTGCTACTCTCTTTGGAGCCCTGATCTTTTACCTCC 6994  
Db 2283 TAGGACCTTACGAGCTAATATTTGCTACTCTCTTTGGAGCCCTGATCTTTTAACTCC 2342  
Qy 6995 TTGTTAACTTTGCTCTTCCAGAACTGTAAACCTTACAAATGGAGCCCAAGATGC 7054  
Db 2343 TTGTTAACTTTGCTCTTCCAGAACTGTAAACCTTACAAATGGAGCCCAAGATGC 2402  
Qy 7055 AGTCCAAGACTAAGATCTACCGCAGACCCCTGGAACCGGCTGTAGCCCAAGATCTGATG 7114  
Db 2403 AGTCCAAGACTAAGATCTACCGCAGACCCCTGGAACCGGCTGTAGCCCAAGATCTGATG 2462  
Qy 7115 TTAATGACATCAAGGACCCCTCTGAGGAAATCTCAGTGTGCAACAACCTTCTACTAGCC 7174  
Db 2463 TTAATGACATCAAGGACCCCTCTGAGGAAATCTCAGTGTGCAACAACCTTCTACTAGCC 2522



Db 1597 TACCTCAAACTCACCTGTGTAAATTTAGCAATCTATAGACACAGCCAACTCCCAATG 1656  
Qy 6333 CATCAGGTGGTAACCTCCCAACACAAATAGTCTGCCTACCTCTAGGAATATTTTGT 6392  
Db 1657 CATCAGGTGGTAACCTCCCAACACAAATAGTCTGCCTACCTCTAGGAATATTTTGT 1716  
Qy 6393 CTGTGTACCTCAGCCCTATCGTTGTTTGAATGCTCTTCAGAACTCTATGCTTCCTCTC 6452  
Db 1717 CTGTGTACCTCAGCCCTATCATTTGTTGATGCTCTTCAGAACTCTATGCTTCCTCTC 1776  
Qy 6453 ATTCTTAGTCCCCVATRCATCTACAGTAAAGATTTATACAGTTATCTCATATC 6512  
Db 1777 ATTCTTAGTCCCCVATRCATCTACAGTAAAGATTTATACAGTTATCTCATATC 1836  
Qy 6513 TAAGCCCGCAACAAAGAGTACCCATCTCTCTTTGTTATAGGAGCAGGAGTGTAGG 6572  
Db 1837 TAAGCCCGCAACAAAGAGTACCCATCTCTCTTTGTTATAGGAGCAGGAGTGTAGG 1896  
Qy 6573 TGCACCTAGTACTGCGCAATTTGGCGGTATCAAACTCTACTCAGTCTACTACAACTATC 6632  
Db 1897 CGGAGTAGTACTGCGCAATTTGGCGGTATCAAACTCTACTCAGTCTACTACAACTATC 1956  
Qy 6633 TCAAGAACTAAATGGGGAATGGAACGGGTGCGGACTCCCTGGTCACTTGCAGATCA 6692  
Db 1957 TCAAGAACTAAATGGTGCATGGAATGGGTGCGGACTCCCTGGTCACTTGCAGATCA 2016  
Qy 6693 ACTTAACCTCTAGCAGCAGTACTCTTCAAACTCAAGAGCTTTAGCTTAACCGC 6752  
Db 2017 ACTTAACCTCTAGCAGCAGTACTCTTCAAACTCAAGAGCTTTAGCTTTGTAACCGC 2076  
Qy 6753 TGARAGAGGGGAACTGTTTATTTTGGGGAAGATGCTGTTATTTGTTAATCAATC 6812  
Db 2077 GGAAGCGGGGAACTTTTATTTTGGGGAAGATGCTGTTATTTGTTAATCAATC 2136  
Qy 6813 CGGAATCGTCACTGAGAAAGTTTAAAGAAATTSAGATCGAATAACAAGTACAGCA 6872  
Db 2137 CGGAATCACTACGAGAAAGTTTAAAGAAATTTCAAGTCTGAATAAAGTACAGCA 2196  
Qy 6873 GCTTCGAAACACTGAGCCCTGGGCTCTCAGCCTATGATGCTGATTCCTCCCTT 6932  
Db 2197 GCTTCGAAACACTGAGCCCTGGGCTCTCAGCCTATGATGCTGATTCCTCCCTT 2256  
Qy 6933 CTTAGGACCTTAGCAGCTATATATTTGCTACTCTCTTTGGACCTGTATCTTTACCT 6992  
Db 2257 CTTAGGACCTTAGCAGCTATATATTTGCTACTCTCTTTGGACCTGTATCTTTACCT 2316  
Qy 6993 CTTGTTAACTTTGCTCTTCCAGATCGAAGCTGTAACTA-----CAAAAT 7040  
Db 2317 CTTGTTAACTTTGCTCTTCCAGATCGAAGCTGTAACTA-----CAAAAT 2376  
Qy 7041 GGAGCCCAAGATGAGTCAAGATCTAAGATCTAAGATCGGAGCCCTGAGCCGCTGTAG 7100  
Db 2377 GGAGCCCAAGATGAGTCAAGATCTAAGATCGGAGCCCTGAGCCGCTGTAG 2436  
Qy 7101 CCCAGATCTGATTTAATGATCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCA 7160  
Db 2437 CCCAGATCTGATTTAATGATCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCA 2496  
Qy 7161 ACCTCTACTACGCCCCAAATTCAGCAGGAAAGCAGTGTAGCGGTGTGCGGCAAGCTCC 7220  
Db 2497 ACCTCTACTACGCCCCAAATTCAGCAGGAAAGCAGTGTAGCGGTGTGCGGCAAGCTCC 2556  
Qy 7221 AACAGCACTTAGTCTTCTGTTAGATGGGCACTGAGAGCAGGACTAGCTGGATTTTC 7280  
Db 2557 AACAGCACTTAGTCTTCTGTTAGATGGGCACTGAGAGCAGGACTAGCTGGATTTTC 2616  
Qy 7281 CTAGGCTGATTAAGATCTTAAGCTAGTGGGAGGTGACCAATCCACCTTTAAACA 7340  
Db 2617 CTAGGCTGATTAAGATCTTAAGCTAGTGGGAGGTGACCAATCCACCTTTAAACA 2676  
Qy 7341 CGGGGCTTCAACTTAGTCACTGACCTGACCAATC-----AGAGACTCACTAAAT 7391  
Db 2677 CGGGGCTTCAACTTAGTCACTGACCTGACCAATC-----AGAGACTCACTAAAT 2736

Qy 7392 GCTAATTAGGCAAGACAGAGGTAAAGAAATAGCCAAATCATATTTATG 7439  
Db 2737 GCTAATTAGGCAAGAAACAGGAGATAAAGAAATAGCCAAATCATCTGTTG 2784

## RESULT 11

US-09-949-016-15858/c  
; Sequence 15858, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15858  
; LENGTH: 15858  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-15858

Query Match 29.6%; Score 2220.4; DB 3; Length 145320;  
Best Local Similarity 86.6%; Pred. No. 0;  
Matches 2728; Conservative 92; Mismatches 216; Indels 115; Gaps 32;

Qy 606 GATGGAGACGTTCCCGCAAGACAGCAAAACGCCCTAAGACGTTATCTGGAATTTGGGA 665  
Db 97137 GATAGAAACGTTCCCTCAAGGCAAAACACCCCTAAGATGTTATCTGGAATTTGGGA 97078  
Qy 666 MCAATTTGACCCCTCAGACACTAAGAAAGAAAGACGTTATCTTCTGACAGTCCGCCCTG 725  
Db 97077 CCAATTTGACTCTCAGATGCTAAGAAAGAAAGAC--ATATTTCTTCTGACGACCGCCTG 97020  
Qy 726 GCA-----CTCTGAGGGAGTATAAATATAA 753  
Db 97019 GCAAGCATATACTCTTTAAGGGGAGAAACCTGGCATCTCTGAGGAAGCATAAATATAA 96960  
Qy 754 CACCATCTACAGCTAGACGTTCTTTTGTAGAA---AAGCAATGAGTGAAGTGCATA 810  
Db 96959 CACCATCTTACAGCTAGACCTCTTTTGTAGAAAGAGGCAAAATGGTGTGAAGTGTCTATA 96900  
Qy 811 AGTACAAACTTTCTTTTCAATTAAGAGACAACTCACAATTTATGTAAAGTGTGATTTATG 870  
Db 96899 CGTACAACTTTCTTTTCAATTAAGAGACAACTCGCAATTTATGTAAAGTGTGATTTATG 96840  
Qy 871 CCCTACAGGAAGCTTCAGAGTCTACCTCTCTATCCAGCATCCCC--GACTCTTTCCCC 928  
Db 96839 CCCTACAGGAAGCTTCAGAGTCTACCTCTCTATCCAGCATCCCCAGACTCTTTCCCC 96780  
Qy 929 AMYTAATAGGACCCCTTCAACCAATGTTCCAAAGGAGATAGACAAAGGTTAAA 988  
Db 96779 AAATAATAGGACCCCTTCAACCAACCGTCCAAAGGAGATAGACAAAGGTTAAA 96720  
Qy 989 CAGTCAACCAAGAGTGCATATTTCCCAATTTATGACCCCTCCCAAGCAGTGGGAGAA 1048  
Db 96719 CAACTAACCAAGAAATGCCAATTTCCCGGATTTATGCCCTCC-----AAGCGGTGGGAG 96664  
Qy 1049 GAGAAATCGGCCAGCAGAGTGCATGCTTTTCTCCAGACTTAAAGCAAAATAA 1108  
Db 96663 GAGAAATCGGCCAGCAGAGTGCATGCTTTTCTCTCTCAGACTT---TAAATTA 96608  
Qy 1109 AAACAGACTTAGGTAATTTCTCAGATAAYCTGTAGTGGCTATATTTGTTTTCACAGGTT 1168









```
Db 30044 CAAGGACACTTTTAAAGAAATGTGTCCAAATAGAAATAAGCCGCCCTCGTCATGCACC 30103
Qy 1946 TTATTTCAAGGAATCACTGGAAGCCCACTGCCCCAGGGACAAGCTCTTTGAGTCA 2005
Db 30104 TCGTGTCAAGGAATCACTGTAAAGCCCACTGCCCCAGGGACGTAGGTCTCTGAGTCA 30163
Qy 2006 GAAGCCACTAAACAGATGATCAGCAGCAGGAGTCTGAGGGTGCTGGGGCAAGCGCATCC 2065
Db 30164 GAAGCCACTAACCAAGATGATCAGCAGCAGGAGTCTGAGGTGCCCGGGCAGCAGCCAGCC 30223
Qy 2066 CATGCCATCACCCTCAGAGAGCCCTGGGTATGCTTGAACATTTGAGGGCAGGAAGGT --- 2122
Db 30224 CATGCCATCACCCTCAGAGAGCCCTGGGTATGCTTGAACATTTGAGGGCAGGAGGCTAAC 30283
Qy 2123 TGCTCTCTGGACACTGGTGGCTCTTCTAGTCTTACTCTTCTGTCGCGGACAACTGTCC 2182
Db 30284 TGCTCTCTGGACACTGGTGGCTCTTCTAGTCTTACTTCTGTCGCGGACAAAGCGTCC 30343
Qy 2183 TCCAGATCTGTCACTATTCTGAGGGGCTCCNTAAGACGGGAGTCACTAGATACCTTTTTC 2242
Db 30344 TCCAGAGCTGTCACTATCC --- AAGGGTCTTAGACAGCCAGTCACTAGATAC -TTCTC 30399
Qy 2243 CCAGCCACTAAGTTATGAACCTGGGAGCTTTATCTTTTTCATGCTTTTCTTAATTATGC 2302
Db 30400 CCAGCCACTAAGTTGTG -ACTGGGAACTTCACTCTTTTTCATGCTTTTCTTAATTATGC 30458
Qy 2303 TTGAAGCCCACTACCTTGTGTAGGAGAGACATTTCTAGCAAAAGCAGGGGCCATTATAC 2362
Db 30459 CTGAAGGCCCACTCCCTTGTGTAGGAGAGACATTTCTAGCAAAAGCAGGGGCCATTATAC 30518
Qy 2363 ACCTGAACATAGGAGAAACACCGTTTGTGTGTCCTCTTGTAGGAGGAAATTAAT 2422
Db 30519 ACCTGAACATAGGA --- GAACACCGTTTGTGTGT -CCCTCTTGTAGGAGGAAATTAAT 30573
Qy 2423 CCTGAAGTCTGGGCAACGAGCAATATGACGAGGCAAAAGATGCCCGTCTCTGTTCA 2482
Db 30574 CTTGAAGACTGGGCAACGAGCAATATGACGAG -CAAGATGCCCGTCTCTGTTCA 30632
Qy 2483 AGTTAAACTAAAGGATTCACATTCCTTCTTCCCTAACCAAGGAGTACCCCTCAGACCCA 2542
Db 30633 AGTTAAACTAAAGGATTCCTGCTCTCTTCCCTCACCACCAAGGAGTACCCCTTAGACCGGA 30692
Qy 2543 GSCCACAAGGATTCCAAAAGATTTGAAGACTTAAAGCCCAAGGCTTAGTAAAC 2602
Db 30693 GGCTCAACAGGACTTCCAAAAGA --- TTAAGGACCTTAAAGCCCAAGGCTTAGTAAAGC 30749
Qy 2603 ATGCATAACTCCCTGCAGTAATTCGCTAGTGTGATTCAGGAGGACAGAAACCCAGTGGAC 2662
Db 30750 ATGCAATAGCCCTACAATAATCC --- AACTTTAGGAGTACAGAAACCCAGTGGAC 30802
Qy 2663 AGTGAGGTTAGTGAAGATCTCAGGATTAATCAATGGAGGCCGTTGTCTTTTATACCC 2722
Db 30803 AGTGA -GGTTAGTGAAGATCTCAGGATTAATCAAT -GAGGTCACTGTCCCTCTATACCT 30860
Qy 2723 AGCTGTACTAGCCCTTACTGTGMYTCCCAATATACAGAGGAGCAGAGTGGTTTAC 2782
Db 30861 AGCTGTACTAGCCCTTATTTCTGCTTTCCCAATATACAGAGGAGCAGAGTGGTTTAC 30920
Qy 2783 ATCTCTGACCTTMAAGATGCTTCTCTGCACTCCCTGTACATCCTGACTCTCAATCTT 2842
Db 30921 AGACCTGACCTTMAAGATGCTTCTCTGCACTCCCTGTACATCCTGACTCTCAATCTT 30980
Qy 2843 GTTTCCTTTTGAAGATCTTCAAAACCCARCATCTCAACTCACTCGACTTTTTTACC 2902
Db 30981 ATTTGCTTTTGAAGATCTTCAAAACCCARCATCTCAACTCACTCGACTTTTTTACC 31040
Qy 2903 AGGGTTCAAGGATAGYCCCATCTATTTGGCCAGGATTTAGCCCAAGCTTGAGTCARTY 2962
Db 31041 AGGGTTCAAGGATAGYCCCATCTATTTGGCCAGGATTTAGCCCAAGCTTGAGCCGGTT 31100
Qy 2963 MTCATACCTGGACACTCTTGTCTCTTCGTAGTGGATGA -TTTACTTTTTCGTCGCT 3020
Db 31101 CTCATACCTGGGCACTCTTGTCTCTTGTGTATGTTGATGATTTTACTTTTAGCCCGCAGT 31160
```

---

```
Qy 3021 TCAGAAACCTTGTGTCATCAAGCCACCACCAAGRCCTCTTMAATTTCTCGCYACCTGTGCG 3080
Db 31161 TCAGAAACCTTGTGTCATCAAGTCACCCACCAAGTCTCTTAAATTTTCTCGCTACCTGTGCG 31220
Qy 3081 TACAWGTTTCCAAACASARAGCTCARTCTGTCTCACAGCAGGTTAAATACTTAGRCCTA 3140
Db 31221 TACAAGTTTCCAAACCAAGGCTCAGCTCTGTCTCACAGCAGGTTAAATACTTAGSGCTA 31280
Qy 3141 ABATTTATCAAGGACACGAGCCCTCAGTGAGGAAATATCCAGCCTATCTAGCTTAT 3200
Db 31281 AAATTTATCAAGGACACGAGGCGCTCAGT -----GCCTATTCTGGCTTAT 31326
Qy 3201 CCTCATCTCAAAACCCCTAAAGCAACTAAGRRRTTCTTGGCRTPAAYAGGVTTCGCCGA 3260
Db 31327 CCTCATCTCAAAACCCCTAAAGCAACTAAGAGATTCCTTGACATTAACAGGTTTCTGCCAA 31386
Qy 3261 AWATGGATTTCCCGAGTGTGCRAAATAGCCAGGTCATTAWATACASTAAATTAAGGAAC 3320
Db 31387 ATATGGATT -CCAGGTACGGGAAATAGCCAGACCATTTATACACTTAATTAAGGAAC 31445
Qy 3321 TCAGAAAGCCAAATACCAATTTARTAAGATGAGAYAMCTGAAGYMRAGTGGCTTTCCAGGC 3380
Db 31446 TCAGAAAGCCAAATACCAATTTAGTAAGATGAGACACCTGAAGCAGAAAGCGCTTTCCAGGC 31505
Qy 3381 CCCTAAAGAGGCGCTTAAACCCAAAGYCCAGTGTTAAGYTTGCCAACGCGGCAAGACTTT 3440
Db 31506 CCTAAAGAGGCGCTT -AAACCCAAAGCGCCCTGTTAGCTTGGCCAAACGGGCGAAGACTTT 31563
Qy 3441 TSTTVAAYRTTCAGAAAAAACAAGAAAYAGTCTTGGAGTCTTACACAGRTCCRAGG 3500
Db 31564 TCTTTACATGTACAGAAAAAACAAGAAATAGCTCTAGAGTCTCTTACACAGGTC - - - 31618
Qy 3501 GAYGAGCTTGCAACCTGTGCRYACCTGAATAAGGAAATGATGTAGTGGCAAGGGTTG 3560
Db 31619 GATGAGCTTGCAACCCATGGCATACCTGAGTAAGGAAATTCATGTAGTGGCAAGGGTTG 31678
Qy 3561 RCYTCATTTTATGCGTGTAGTGGCGAGTGGAGTACAGTCTAGTATCTGAAGCAGTTAAAT 3620
Db 31679 GCCTCATTTTATGCGTGTAGTGGCGAGTGGAGTCTAGTATCTGAAGCAGTTAAAT 31738
Qy 3621 AATCAGGAGAGATCTTACTGTGTGACATCTCATCAKGTGAAYRCATACCTCACTGC 3680
Db 31739 AATCAGGAGAGAT - - - -CTGTGTAGACATCTCATCAKGTGAAGCAGTACTCACTGC 31794
Qy 3681 TAAAGGAGACTTGTGGCTGTGCAGAACACT 3711
Db 31795 TAAAGGAGACTTGTGGCTGTGCAGAACACT 31825

RESULT 13
US-09-949-016-5640
; Sequence 5640, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5640
; LENGTH: 2763
; TYPE: DNA
; ORGANISM: Human
```

US-09-949-016-5640

```
Query Match      29.4%; Score 2204; DB 3; Length 2763;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 2194; Conservative 35; Mismatches 0; Indels 1
```

Qy	5352	GATACATCCTGGGAAGGACCCTACCCAGTCAATTTTATVTAACCCCAACTGCGGTAAAAGTG	5411
Db	535	GATACATCCTGGGAAGGACCCTACCCAGTCAATTTTATCTACCCCAAAGTCGCGTTAAAAGTG	594
Qy	5412	GCTGGAGTGGAGTCTTGATATACATCACACTTGAGTCAAACTCTCGATACATGCCAAGGAA	5471
Db	595	GCTGGAGTGGAGTCTTGATATACATCACACTTGAGTCAAACTCTCGATACATGCCAAGGAA	654
Qy	5472	CCTGAAATCCAGGAGAACCGCTAGCTATTCTGTGAACCTCTAGAGGATTTGCGCCTG	5531
Db	655	CCTGAAATCCAGGAGAACCGCTAGCTATTCTGTGAACCTCTAGAGGATTTGCGCCTG	714
Qy	5532	CTCTTCAACAACCAACGAGGAGAAAGTAACTAAATCATAAATCCCATGSGCCTCC	5591
Db	715	CTCTTCAACAACCAACGAGGAGAAAGTAACTAAATCATAAATCCCATGSGCCTCC	773
Qy	5592	TTATCATATTTTTCTCTKTASTGTTSTTVYACCCSTTTTCACTCTCACTGCAACCCCTCC	5651
Db	774	TTATCATATTTTTCTCTKTASTGTTSTTVYACCCSTTTTCACTCTCACTGCAACCCCTCC	833
Qy	5652	ATGCCGCTGTATGACAGTAGTCTCCCTCYACMAGAGTTTCTATGGAGAATCGACGCTCC	5711
Db	834	ATGCCGCTGTATGACAGTAGTCTCCCTCYACMAGAGTTTCTATGGAGAATCGACGCTCC	893
Qy	5712	CGGAAATATTGATGSCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCACTTCACATGC	5771
Db	894	CGGAAATATTGATGSCCCCATCGTATAGGAGTCTTTSTAAGGGAACCCCACTTCACATGC	953
Qy	5772	CCACACCATAATGCCCGCCAACTGCTATCACTCTGCCACTCTTTGGCATGATGCAAAATAC	5831
Db	954	CCACACCATAATGCCCGCCAACTGCTATCACTCTGCCACTCTTTGGCATGATGCAAAATAC	1013
Qy	5832	TCATTATTGACAGGAAAAATGATTAATCTAGTTGTCCTGGAGGAATTGGAGTCACTGT	5891
Db	1014	TCATTATTGACAGGAAAAATGATTAATCTAGTTGTCCTGGAGGAATTGGAGTCACTGT	1073
Qy	5892	CTGTTGACTTACTTTCAACCCAAACTGCTATGTCGTAGTGGGGGTGGAGTTCAGATCAGGC	5951
Db	1074	CTGTTGACTTACTTTCAACCCAAACTGCTATGTCGTAGTGGGGGTGGAGTTCAGATCAGGC	1133
Qy	5952	AAGAAAAACATGTAAAGAAGTAATCTCCCAACTCACCGGGTACATGGCACCTCTAG	6011
Db	1134	AAGAAAAACATGTAAAGAAGTAATCTCCCAACTCACCGGGTACATGGCACCTCTAG	1193
Qy	6012	CCCTACAAAGGACTAGATCTCAAAACTACATGAACCCCTCCGTAACCATCTCGCCT	6071
Db	1194	CCCTACAAAGGACTAGATCTCAAAACTACATGAACCCCTCCGTAACCATCTCGCCT	1253
Qy	6072	GGTAAGCCTATTTAATACCAACCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCCTAC	6131
Db	1254	GGTAAGCCTATTTAATACCAACCTCACTGGGCTCCATGAGGTCTCGGCCCAAAACCCCTAC	1313
Qy	6132	TAACTGTTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATCCTCTGACCTGA	6191
Db	1314	TAACTGTTGGATATGCTCCCTCCCTGAACTTCARGCCATATGTTTCAATCCTCTGACCTGA	1373
Qy	6192	ACAATGAAACACTTCAGCAGAGNAATAAACACCACTTCGGTTTTAGTAGGACCTCTTGT	6251
Db	1374	ACAATGAAACACTTCAGCAGAGNAATAAACACCACTTCGGTTTTAGTAGGACCTCTTGT	1433
Qy	6252	TTCCAATSTGGAATAAACCCATACCTCAAACTCACCTGTGTAAATTTTAGCAATACTAC	6311
Db	1434	TTCCAATSTGGAATAAACCCATACCTCAAACTCACCTGTGTAAATTTTAGCAATACTAC	1493
Qy	6312	ATACAAACCAACTCCCAATGCAATCAGGTGGGTAACTCTCTCCCAACAAATAGTCTGCCT	6371
Db	1494	ATACAAACCAACTCCCAATGCAATCAGGTGGGTAACTCTCTCCCAACAAATAGTCTGCCT	1553

```

Qy 7452 GCAGGAGGGACAATGATCGGATATAAACCCAAAGTCTTCAGCGGCAACGGCAACCCCC 7511
Db 2634 GCAGGAGGGACAATGATCGGATATAAACCCAAAGTCTTCAGCGGCAACGGCAACCCCC 2693
Qy 7512 TTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTTCACCTCTATTAAATC 7571
Db 2694 TTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTTCACCTCTATTAAATC 2753
Qy 7572 TTGCARCTGC 7581
Db 2754 TTGCNACTGC 2763

RESULT 14
US-09-573-080A-26
; Sequence 26, Application US/09573080A
; Patent No. 6828097
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL
; APPLICANT: ROGAN, PETER
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/573,080A
; CURRENT FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 8399
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat region
; LOCATION: (1)..(8399)
; OTHER INFORMATION: herv9
; PUBLICATION INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: Jurka, J; Walichiewicz, J; Milosavljevic, A
; TITLE: Prototypic sequences for human repetitive DNA
; JOURNAL: Journal of Molecular Evolution
; VOLUME: 35
; ISSUE: 4
; PAGES: 286-291
; DATE: 1992-10-
; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
; DATABASE ENTRY DATE: 1996-01-26
; DATABASE ENTRY DATE: 1996-01-26
US-09-573-080A-26

Query Match 28.9%; Score 2168; DB 3; Length 8399;
Best Local Similarity 67.6%; Pred. No. 0;
Matches 3334; Conservative 143; Mismatches 1354; Indels 103; Gaps 34;

Qy 605 TGATGGGAAACGTTCCCGCGCAAGACAAAAAGCCCGCTAAGACGCTATTCTCGGAAATTTGGG 664
Db 1276 TGATAGGAAACACTCAGCATCAACACGCTCACCTTTGAAATGTATCCTAAGCCATTGGG 1335
Qy 665 AMCAATTTGACCTTCAGACACTAAGAAGAAACGACTTATATCTTCGAGTGGCCGCT 724
Db 1336 ACCAATTTGACCCCAAAACCCCTGAAAAGAGGCGACTCTTTTCTGCACACTACAGCT 1395
Qy 725 GGCACT-----CCTGAGGGAAGTATAAATTATA 752
Db 1396 GACCTAATATTCTCTCTGATGGGNAAAATGGCCACCTTGAGGGAAGTATAAATTACA 1455
Qy 753 ACACCATCTTACAGCTAGACTCTTTTGT---AGAAAAGGCAAAATGAGTGAAATGCGCAT 809
Db 1456 ATACTACTCTCAGGTTGACCTTTCTGTAAGAGGGAAGGTAAATGGAGTGAATACCAT 1515
Qy 810 AAG----TACAACTTTCTTTTCATTAAAGACAACTCACAAATTTATGTTAAAGTGTGAT 865
Db 1516 AAGGTAATTACAAAGCTTTCTTTTCTACTGAAGGAGAATAACAACTATGCAAGCTTGAAT 1575
Qy 866 TTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCAGCACTCCCGACTCCTTC 925

```

Db 2641 CTACTATGCCGAGCAANTCACTGGAAGGTGCACTGCCCCAGAGACAAAAGTTTCTTAGG 2700  
Qy 2003 TCAGAGCCACTAACAGATGATCCAGACAGAGACTGAGGGTGCCTGGGGCAAGCCCA 2062  
Db 2701 TTGGAAGCCCACTGATGATCCAAACAGGACTGAGGGTGCCTGGGGCAAGCCCA 2760  
Qy 2063 TCCCATGCCATACCCCTCACAGAGCCCTGGGTATGCTTGACCAATGAGGGCCAGGAA 2119  
Db 2761 GCTCATGTCACTCCTGAGCCCAAGTATGTTTAAACATGAGGGCCAGGAAAT 2820  
Qy 2120 GGTGTCTCTGAGCACTGGTGGGTCTTCTAGTCTTACTCTTCTGTCGCCGACAACTG 2179  
Db 2821 GACTTCTCTGGACACTGGCAGCTCTCTCAGTGTCTATCTCTTCTGATGACTG 2880  
Qy 2180 TCTTCAGATCTGTCACTATCTGAGGGGTCCNTAAGAGGGGAGTCACTAGACTTTT 2239  
Db 2881 TCTCAGGTCTGTGTACATCC---GAGGAATCCTGGGGCAGCGCTGTAACAGCT-GTTT 2936  
Qy 2240 YTCAGCCACTAAGTTATGAACTGGGGAGCTTATCTTTTACATGCTTTCTAATTA 2299  
Db 2937 CTCCACCTCTCAGTTGT-AATGGGAGACTTGTCTTTTCTGTATGCTTTCTTGTGA 2995  
Qy 2300 TGTGTGAAGCCCACTACCTTGTGTAGGAGAGACATTTAGCAAAAGCAGGGGCCATTA 2359  
Db 2996 TGCTGAAAGTCCACACCTTTATTAGGAGATATAT--TAGCCAAAGCTGGAGCTATTA 3053  
Qy 2360 TACACCTGAACATAGGAAAGGAACACCCGTTTGTGTGTCCTGCTGTGAGGAAGAAAT 2419  
Db 3054 TCTACATGAATATGGGAAACAAGTTACCCATTTTGTGTCCTTCTGTATGCTTTCTTGTGA 3112  
Qy 2420 AATCCTGAAGTCTGGGCAACAGAGGACATATGGAGGACCAAGATGCCGCTCTGT 2479  
Db 3113 AACCTGAAGTCTGGGCAATTTGGAAGGACAAATTTGGAAGGGCAAAA-AATGCCACCCAGT 3171  
Qy 2480 TCAAGTTAAACTAAAGGATCCACTTCTTCTTCCCTACCAAGGAGTACCCCTCAGACC 2539  
Db 3172 CCAATCAGCTAAAGACTCCACCACTTCTTCTTATCAAAAGCAATATCCCTTAAGGCC 3231  
Qy 2540 CAAGGCCCAACAAAGGATTCAAAGATTTGTAAGGACTTAAAGCCCAAGGCTTAGTAA 2599  
Db 3232 TGAAGCTCTTAAAGGATACAGGATATTTGTAACATTTTAAAGCTCAAGGCTTAGTAG 3291  
Qy 2600 ACCATGATTAATCCCTGAGTAAATTCCTGAGTGGATGAGGAGGACAGAAACCCAGT 2659  
Db 3292 GAAATGAGCAGTCCCTGCAACCC---AATTTAGGAGTACAAAAACCAATG 3344  
Qy 2660 GACAGTGGAGGTAGTGAAGATCTCAGGATTTCAATGGAGGCGCTTGTCTTTTATA 2719  
Db 3345 GTCACCTGGA-GACTAGTGAAGATCTTAGACTCATCAAT-GAGGCAATTTCTCTATA 3402  
Qy 2720 CCAGCTGTACCTAGCCCTTATCTGTGMYTTCCCAATACCAAGAGAAAGCAGTGGTT 2779  
Db 3403 TCCAGTTGTACCAACCCCTATACCTGCTTTCTCAANTACAGAGAGAGCAATGGTT 3462  
Qy 2780 TACASTCTGGACTTMAAGATGCTTCTGCTGATCCCTGATACATCTGACTCTCAAT 2839  
Db 3463 CATGGTCTGGACCTCAAGGATGCTTCTGCTTCCC-CTGCACTCTGACTCCAGTT 3521  
Qy 2840 CTGTTTGTGCTTGAAGATCTTCAACCCACATCTCAACTCCCTGGACTTTTACC 2899  
Db 3522 TCTGTTTGTGCTTGAAGATCCCAAGACCAACAGTCCCACTTACATGATGCTGTGCC 3581  
Qy 2900 CCAAGGTTTCAGGGATAGYCCCATCTATTGGCCAGGCAATAGCCCAAGACTTGAGYCA 2959  
Db 3582 ACAAGGTTTAGGGATAGCCCTCACTGTTGTGTGAGGACATGCGCCCAAGATCAGGCCA 3641  
Qy 2960 RTWTCACTCGACACTTCTGCTCTTCTGTAGTGGATGATTTACTTTTTCYCCYR 3019  
Db 3642 CTCTCAAGTCCAGGCACCTTGTGCTCTCAGTATGTGGATGATTTACTTTTGGCTACCAG 3701  
Qy 3020 TTCAGAAACCTTGTGCCATCAAGCCCAAGRCCTTAAATTTCTGCTGCTGCTGCTGG 3079  
Db 3702 TTCAGAAAGCCTCATGCCAGCAGGCTACTCTAGATCTCTTGAATCTTCTACCTAATCAAG 3761

Qy 3080 CTACAWGGTTTCAAAACSAARAGCTCARTCTCTCACAGCAGGTTAAATACTTAGGRC 3139  
Db 3762 GTACAAGGCATCTAGGTCAAGGTGCGAGCTTCTCTACAGCAGGCTAAATATCTAGGCT 3821  
Qy 3140 ABAATATCCAAAGGCACCAAGGCTCAGTGAAGAAATATCCAGCTTACTAGGCTTA 3199  
Db 3822 AATCTTAGCAGAGGACCAAGGCTCAGCAAGGAATGAATACAGCTTACTAGGCTTA 3881  
Qy 3200 TCTCATCTCAAAACCTTAAAGCAACTAAGRRRTTCTTGGCTTAAYAGTTCCTGCG 3259  
Db 3882 TCTTCCCTTAAGACATTAACAGTTGAGGGTCTTGGAAATCACCGCTTTTGGCG 3941  
Qy 3260 AAWATGATTCCTCCAGGTGTCRAAATAGCCAGGYCATTAATACASTAATTAAGGAAA 3319  
Db 3942 ACTATGGA-TCCCTGATACAGTACAGTACAGGCTTCCCATCTCTAATCAAGGAGA 4000  
Qy 3320 CTCAGAAAGCAATACCCATTTARTAAGATGGAYANCTGAAGYMAAGTGGCTTTCAGG 3379  
Db 4001 CCCAGAGGCAATATCTTATCTAGTAGAATGGTAAACCAAGGGCAGAAACAGGCTTCAAAA 4060  
Qy 3380 CCCTTAAAGAGGCTTAAACCCCAAGYCCAGTGTAAAGYTTGCCCAACRGGGCAAGACTT 3439  
Db 4061 CCTTAAAGCAGGCTTGA---CAAGTTCAGCTTTAAGCTTCCAC-----AACTT 4111  
Qy 3440 TTSYATATYRTCACAGAAAAAACAAGAAAGTCTTGGAGTCTTACACAGTCCRAG 3499  
Db 4112 CTCTTATATGTACAG-AGAGAGCAGGATAGTCTTGGAGTCTTACTCAGACTCTTG 4170  
Qy 3500 GGNAGAGCTTGCACCCYRTGGCYACCTGASTAAGGAAATATGATGTAGTGGCAAGGTT 3559  
Db 4171 GGCAACCCCAACACAGTGGCATCTAAGTAAGGAAATTTAGTAGTACGCAAAAGCT 4230  
Qy 3560 GRCYCATTTTAYGGGTAGTGGTGGAGTACAGTGTAGTATCTGAAGCAGTTTAAAA 3619  
Db 4231 GGGCTCATGTTTAAAGGTAGTGGAGTGGCACTGGCCATCTTCATGTGAGAGGCTTATCAAAA 4290  
Qy 3620 TAATACAGGAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATACTCAGT 3679  
Db 4291 TAATACAGGAA-AGATCTCAGTCTGCTGACTACTCATGTTGTAATGCACTAGGTG 4349  
Qy 3680 CTAAAGGAGACTTGGCTGTGAGCAACAGTGTCTTAAATTCAGGCTCTATTTACTTG 3739  
Db 4350 CCAAGGAAATTTATGGCTATCAGACTACACCTACTAGTACACGAGGACTACTCTTG 4409  
Qy 3740 AARGGCACTGTGCRACGTGCACTTGTGCACTTAAACCCAGYCNCACTTCTTCCAG 3799  
Db 4410 AGGACCACTGCTTCAATATCTATGTTGGGCTCAACCTGCGCACTTTTCTCCAG 4469  
Qy 3800 ACATGAAGAAAGATARAAYATACTGTCAACAAATAATTTCTCAAACTATGCCACTC 3859  
Db 4470 AGAATGGGNAACCAATGAGCATGACTGCCAACAAATACAGTCCAGACTTATGCCACCC 4529  
Qy 3860 GAGGGAGCTTGTAGAGTTCCTGAGCTGATCCYAGCCTTCAACTTGTATCTGATGGA 3919  
Db 4530 AAGATGATCTCTAGAGTCCCTTAGCTTAATCTCTGACCTT-AACTATATACCGATGGA 4588  
Qy 3920 AGTCTCTTGTAGAAAAGGACTTCGAAAAGYGGGTATGCAAGTGTGATGATATGGA 3979  
Db 4589 AGTTCATTTGCAAGAAATGGGATATGAAGGGCAGGTTGTGACATAGTTAGTGTAACT 4648  
Qy 3980 ATATYTTGAAAGTAAATCCCTCACTCCAGGAACTAGTGTCTGTCAGAACTAATAGCC 4039  
Db 4649 GTACCTGAAAGTAAAGCTCTTCCCAAGGACAGCACCCAGTTAAACAGAACTAGTGGCA 4708  
Qy 4040 YTCATYKGGGCACTAGAATTAGGAGAAAGAAAGGGYAAATATATATATACAGACTTRAR 4099  
Db 4709 CTTACTGAGCCTTAGAACTGGGAAAGGAAAGAAATAAATGTGTATACAGATCAAG 4768  
Qy 4100 TATGCTYACCTAGTCNTCCATGCCCATGMRGCAATATGSAAGAAAGGAAATTCCTAAT 4159  
Db 4769 TATGCTTATCTAATCCCTCATGCCATGCTGCAATATGGAAGAAAGGAGGTTCCTAAC 4828



Qy	4160	TCYAGRGAAACCTATCAMAACATCAGGAAGCATTAGGARATATTATYTGCGWGTACAG	4219
Db	4829	TCGGGGTACCCCAATTAATGCCAAGAGAGATTATGAGTATTTGTCATGCATCGCAA	4888
Qy	4220	AAACCTARAGGCTGGMAGTCTTACACTGCVGGGTCTCATANAAGGNAAGRAAAGGAA	4279
Db	4889	AAGCCCAAGGAGGTGGCAGTCTTACCTGCCAAGCCCATCAGAAGGTGAAGGAGAAAG	4948
Qy	4280	ATASAGRGAAATGCCAAGCAKATATTGAAGCMAAAGAGCTGCAAGCGCAGGACCTCCA	4339
Db	4949	GCAGAGAARACCTCAGGCAGATGCTGAGGCCAANAATCTGCCAGTGGATACTCCCA	5008
Qy	4340	TTAGAAATGCTATTAAACTTCCCTTAGTATAGGGTAAATCCCTTCGGGGAACCAAGCCC	4399
Db	5009	TTAGAAATACCTATGGAAGCACCTTGGAATGGAAACAAACCCCTCCAAGAGATTAAGCCC	5068
Qy	4400	CAGTACTCAGCAGGAGAAAACAGATGGGGAACCTCAGCAGG---CAGTTTTCCTCCCTCG	4456
Db	5069	CAGTATTTCCCAATGAAACAGATGGGGAACCTCATGCGGGGATAGTTTCTTCCCTCCA	5128
Qy	4457	GGACGGTTTAGCCACTGAAGAAGGGAATACTTTTGCTCTCAACTCCTCAATGGAATA	4516
Db	5129	GGGTGGTTAAGCAAAAAGAAAGAAAGTACTTATACCGAAGCCAGTCAGTGGAAATA	5188
Qy	4517	CTTAAAAACCTTCATCAAACTTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAA	4576
Db	5189	CTTAAAAACCTCCACCAAACTTTTTCATATGGGTATTGAAAACACTCATCAATGGCCACA	5248
Qy	4577	TCATTATTTACTGAGCAGGCTTTTCAAACCTATCAAGCARATATKTAGGGGCTGTGAA	4636
Db	5249	TCTCTATTTCAAGCGCCAAATCTACTCTGGGCCCTTCCACAGGTAGTCAAAACCTGTGAG	5308
Qy	4637	KTGTGCCARAAATAATCCCTGCCTTATCGCCAACTCTTCAGAGARAAACAAARAACA	4696
Db	5309	GTGTGCCAAGGAATAATCCCTTGCTCATTTGAAGGCCCATTTGGGGGAACAAGAATA	5368
Qy	4697	GGCCATTACCTGRRARAARACTGGCAACTGATTTTACCCCAA--GCCAAACCTCAGGG	4754
Db	5369	GGTCATCTCTGGAGAGGACTGGCAGTTAGACTTACCCCATATGGCCTTAATCAAGGG	5428
Qy	4755	ATTTCAGTATCTACTAGTCTGGGTARATATCTTTCACGGGTTGGGCARAGGCTTCCCTG	4814
Db	5429	ATTTCAATATCTGTGTGGTCTGTGTGTATACCTTTACAAATTTGGATAGAAGCTTCCCTG	5488
Qy	4815	TAGACAGAAAAGGCCCAAGAGTAATAAGGCACCTAGTTCATGAATAATNTCCAGATT	4874
Db	5489	CAAGACAGAGAAGACTCAGGAAGTGATGAAGTCTTAATGATGAATAATNTCTAGATT	5548
Qy	4875	CGGACTTCCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCCAGGCCACAGTAACCCA	4934
Db	5549	TGGGCTTCCCTAAGGCTTACAGAGTGACATGGTCCAG-TTTTAAAGCCACATNACTCA	5607
Qy	4935	GGGAGTATCCAGCGGTTAGGTATACGATATCACTTACCTGCGCTTGAAGGCCACAGTC	4994
Db	5608	GGGAAATGTCAGGCTGTAGGATACAGTATCACCTTCTACTGCGCCCGAGGCCACAATC	5667
Qy	4995	CTCAGGNAGGTCAGAAAATGANTGAAAYACTCAAAGGCATCTTAAAAAGCAACCCA	5054
Db	5668	CTCAGGGNAGGTCAGAAAGGCAAAATGAAACAACCTCGAGGCATTAAGGAATACTAACCAAG	5727
Qy	5055	GGAAACCCACTCACATGGGCTGTCTGTGCTCATAGCCTTAAAAAGAAATCTCGCAACTT	5114
Db	5728	GAA---CATCTCCATGGCCTACTCTTTTGCCCATGGCCTTGTGTAGAAATCCGAATTC	5783
Qy	5115	TCCCCAAAAGCAGGACTTAGCCCATCGAAATGCTGTATGGAAGGCCCTTTCATAACCAA	5174
Db	5784	TCCTCACAAAATGGGGCTCAGTCCATATGAAATGCTGTATGGACAAATTTCTTCAAAA	5843
Qy	5175	TGACCTTGTGCTTGACCCACAGACAG-CCAACCTTAGTTGCAGACATCACTCTTAGCCAA	5233
Db	5844	TGACCTCTCTTGTATGAGGAAAGACAAACTTGTTCAAAGATATATCTTCTTCGGGCAA	5903
Qy	5234	ATATCAACAGGTTCTTAAAAATTAACAGGAACCTATCCCTTGAGAAGGGAAGAAAGAACT	5293

Db	5904	ATATCAACAAAAACCCCTAAAAAACCTACCTGAAGGATGTACAGAGAAAAAGGGAACACAGATT	5963
Qy	5294	ATTCCACCCWGTGCATCGTATTAGTCAAGTCCCTTCYCTCTAAATCCCATCCTCCTAGA	5353
Db	5964	GTTTCAACAGGAGATCTAGTGTGGTCAATCTCTCCCTCTACTCCCATCTATGA	6023
Qy	5354	TACATCCTGGGAAGGACCTACCCAGTCAATTTATVTAACCCCACTGCGGTTAAAGTGGC	5413
Db	6024	CTCTGTGGGAAGGGTCATCTCAATAATCCTCTACCCACACTGCAGTTAAGGTGGT	6083
Qy	5414	TGAGTGGAGTCTTGGATACATACACACTTGAGTCAATCCTGGATACGCCAAAGGAACC	5473
Db	6084	AGGAGTGAATCTTGGATTCAACACACCCGAGTTAAATTTTGGACATCCCTTGAGGAACC	6143
Qy	5474	TGAAATCCAGGAG	5487
Db	6144	TGTGAGACCATCAG	6157
RESULT 15			
US-09-854-867-26			
; Sequence 26, Application US/09854867			
; Patent No. 7014997			
; GENERAL INFORMATION:			
; APPLICANT: JOAN, KNOLL H			
; APPLICANT: ROGAN, PETER K			
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GEN			
; FILE REFERENCE: 30307			
; CURRENT APPLICATION NUMBER: US/09/854,867			
; CURRENT FILING DATE: 2003-05-08			
; NUMBER OF SEQ ID NOS: 613			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 26			
; LENGTH: 8399			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: repeat region			
; LOCATION: (1)..(8399)			
; OTHER INFORMATION: herv9			
US-09-854-867-26			
Query Match			
Best Local Similarity			
Matches 3334; Conservative 143; Mismatches 1354; Indels 103; Gaps 34;			
Qy	605	TGATGGAAACGTTCCCGGACAGACAAACCGCCCTTAAGACGTATTTCTGGARAATTGGG	664
Db	1276	TGATAGMAACACTCAGGCATCAACACGCTCACCCCTTGAATGTATCTTAAGCCATTGGG	1335
Qy	665	AMCAATTTGACCTTCAGACACTAAGAAAGACGACTATATTTCTTCGAGTGGCGCT	724
Db	1336	ACCAATTTGACCCACAAACCCCTGAAAAAGGAGCGACTATTTTCTGCATACAGCT	1395
Qy	725	GGCACT-----CCTGAGGGAAGTATAAATTATA	752
Db	1396	GACCTTAATTTCTCTCTGATGGGAAAAATGGCCACCTGAGGGAAGTATAAATTACA	1455
Qy	753	ACACCATTTACAGCTAGACGTCCTTTTGT---AGAAAAAGCAAAATGAGGTGAAGTGCCAT	809
Db	1456	ATACTATCTTCGAGTTGACCTTTCTGTAAGAGGGAAGTAAATGAGTGAATACCAT	1515
Qy	810	AAG----TACAACTTCCTTCATTAAAGACCACTCACAATTATGTAATAAAGTGTGAT	865
Db	1516	AAGGTATTACAGCTTTCTTTTCACTGAAGGAGATACACAACTTATGCAAAAGCTTGCAAT	1575
Qy	866	TTATGCCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCAGCACTCCCGACTCCTTC	925
Db	1576	TTACATCCACAGGAGACCTTTTCAGTTACCCCATATCTTAGCCTCCCTAGAGCTCTC	1635
Qy	926	C--CCAMTTAATAGGAGCCCCCTTCAACCCAAATGGTCCAAAGGAGATAGACAAAAGG	983



Db 1636 CTTCTATTAAATGGCAAGCCTCTCCAACTCTCCCTGCCAGAGGAATAAGCAAGAA 1695  
 Qy 984 GTAAACAGTGAACCAAGAGTGCATAATATCCCAATATATGACCCCTCCCAAGAGTGGG 1043  
 Db 1696 ATCTCAAGAGGACCAAAACCCCTGGGCTATCGGTTATGTGCCCTT-CAAGCTGTAGG 1754  
 Qy 1044 AGGAAGAGAAATCGGCCAGCAGAGTGCATGTCVTTTTYYTCTCCAGACTTAAAGCA 1103  
 Db 1755 GGGTGGAGAAATTTGGCCCAACCCAGGTACATGTCTCCCTCTGTGATTTAAAGCA 1814  
 Qy 1104 AATAAAAAACAGACTTAGGTAAATTTCTCAGATAAATCTGATGGGTATATATGTTGTTTACA 1163  
 Db 1815 GATCAAGGCAGACTGGGGGAAGTTTTCAGATGATCTCTGATAGGCATAGATGTCTCCA 1874  
 Qy 1164 AGGGTTAGGCAATTTCTTGATCTGACATGGAGATATATATGTCATCTGCTAAATCAGA 1223  
 Db 1875 GGGTCTAGGGCAACCTTCGATCTCACTTGGAGAGTGT-CATGCTATTTGATGATCAAA 1933  
 Qy 1224 CACTAAACCCCAATGAGAGAGTGCACCACTACTGAGCTGAGGTGTTGGCGATCTCT 1283  
 Db 1934 CCTGGCCCTTAATGAAGAATGTGGCTGTAGCTGAGCCTGAGATTTGGAGATACCT 1993  
 Qy 1284 GGTATCTCAGTCAAGTCAATGGATANGGATGACACAGAGGAAGAAATGATTTCCCA 1343  
 Db 1994 GGTATCTTAGTCAAGTAAATG---ATAGATGACAGCCGGAAGAGATAAATTCCTTA 2050  
 Qy 1344 CAGGCCAGCARGCAGTTCAGCTTASACCTCATTTGGGACACAGAAATCAGTAAATCAG 1403  
 Db 2051 CTGGCCAGCAAGCATCCCGCAGTATGGATCCCACTGGGACCTC---AATCAGATCAT 2106  
 Qy 1404 GGAGATTTGTGCTGACAGATTTGCTAACTTGTGTCTASAGGACTAAGGAATACTASG 1463  
 Db 2107 GGGGACTGAGTCTGAACATCTGTTGACCTGTGTTCTGAAGGACTAAGGAATTAGG 2166  
 Qy 1464 AAGAAATCTAAGAAATTAATCAATGATGTCACCATAAACACAGGGGAAGGAAGAAATC 1523  
 Db 2167 AA-AAAGCCCATGATTTATCAATGATGTCACCATTAATCA-GGGAAGGAAGAAATC 2224  
 Qy 1524 CTATGCTCTTTCTGGAGAGACTAAGGGAGGCAATGAGGAAGCGTGCCTCTCTGTACCTG 1583  
 Db 2225 CTTCTGCTCTTTTCAGGTGGCTTCAGGAGACCTTAAAGAAATATATCTCCACTGTCTAC 2284  
 Qy 1584 ACTCTTCTAGGCCCAACTAATCTTAAGGTAAGTTTATCACTAGTCACTGCTGACAGACA 1643  
 Db 2285 AATCACTCAGGGGTCAATGATTTCTAANAAGATAAGTTTATTAACCAATCAGCAGCAGATA 2344  
 Qy 1644 TTAG-AAAAAATCTCAAAAGTCTGCGTAGGCCCGGAGCAAAATCTTAGAAACCTATTGA 1702  
 Db 2345 TCAGTAGAAGCTCCAAAGCAAGCCCTGGGCCCTGNAACAAATATTTAGGCAATTATTA 2404  
 Qy 1703 ACTTGGCAACYTGGTTTTTTTATATAGAGATCAGGAGGAGCAGGGGAAACAGGACAAAC 1762  
 Db 2405 ACTTGGCAACCGTGTGTTCTATATATAGGACCAAGAGGAAACAGGCCCAA---AGGAAA 2460  
 Qy 1763 GGGATTAATAAAGGCAACCGTTTATGATGACCCCTCAGGCAAGTGGACTTTGGAGGC 1822  
 Db 2461 AGAGATTCAGAAAAGGCGCAGCCTTAGTCATGGCCCTCAGACCAACAAACTTTGGTGT 2520  
 Qy 1823 TCTGGAAGGGAAGGAGCTGGCAATTTGATCTTAATAGGCTTGTCTTCCAGTCCGCT 1882  
 Db 2521 TCAGAGGAGACAGAAATGAGAGGCGCAATCAGCTGGTAGGGCTTTGTTATCAGTGTGT 2580  
 Qy 1883 CTACAGGACACTTTTAAAAAAGATTTGTCAGAGTAAAGTAAGCGGCCCTTGGTCCATGC 1942  
 Db 2581 TTACAGGACACTTTTAAAAAAGATTTGTCAGTGAAGAAACAGCTGCCCCCTTATCTGTGT 2640  
 Qy 1943 CCTTATTTCAAGGGAATCACTGGAAAGGCCCACTGCCCCAGGGGCAAAAGGCTTTTGTAG 2002  
 Db 2641 CTACTATGCGAGGCAATCACTGGAAAGGTGCACTGCCCCAGAGGACAAAGTTTCTTAGG 2700  
 Qy 2003 TCAGAGGCCACTTAAACAGATGATCCAGACGAGGACTGAGGCTGCTGGGGCAAGGCCA 2062  
 Db 2701 TTGGAAGCCCCCAACCTGATGATCCAAACACAGGACTGAGGGGTGCCAGGGCAAGGCCA 2760

Qy 2063 TCCATGTCATACACCTCAAGAGCCCTGGGTATGCTTGAACCATTCAGGCCACAGGAA--- 2119  
 Db 2761 GCTCATGTGTCATCACCTCACTGAGCCCAAGGTATGTTAAACCATTCAGGCCACAGGAAAT 2820  
 Qy 2120 GGTGTCTCTCGGACACTGGTGGGTCTTCTTATGCTTCTTCTGCTCCCGGACACTG 2179  
 Db 2821 GACTTCTCTGAGCACTGGACGTCTTCTCAGTGTCTTCTCAGTGTCTTCTTCTGATGACTG 2880  
 Qy 2180 TCTCCAGATCTGTCACTATTCTGAGGGGTCCNTAAGAGCGGCGAGTCACTAGATACTTT 2239  
 Db 2881 TCTCAAGGTCTGTTTACCATCC---GAGGAATCTCTGGGACGCTGTAAACAGT-GTTT 2936  
 Qy 2240 YTCCAGGACACTAAGTTATGAACCTGGGAGCTTTATTTCTTTTCACTGCTTTTCTAATTA 2299  
 Db 2937 CTCCCACTCTCAGTTGT-AAATTGGGAGACTTTGCTCTTCTGTATGCTTTCTTGTTA 2995  
 Qy 2300 TGTCTGAAGCCCACTACTCTGTTAGGGAGGACATTTAGCAAAAGCAGGCGGCATTA 2359  
 Db 2996 TGTCTGAAGTCCACACCTTTTATAGGGAGATATAT--TAGCCAAAGCTGGAGCTATT 3053  
 Qy 2360 TACCTGAACATAGGAGAGGAAACACCCGTTTGTGTGTTCCCTGCTTGGAGAGCAAT 2419  
 Db 3054 TCTACATGATATGGGACAAAGTTTACCATTTGTTGT-CCCTTACTTGGAGGGAAT 3112  
 Qy 2420 AATCTGAAGTCTGGGCAACAGAGCAATATGAGCGAGCCAAAGAAATGCCCGTCTGT 2479  
 Db 3113 AACCCCTGAAGTCTGGGCAATTTGGAAGGACAAATTTGGAAGGCAAAA-AAATGCCACCCAGT 3171  
 Qy 2480 TCAAGTTAAACTAAGGATTCACCTCTCTTCCCTACCAAGGAGTACCCCTCAGACC 2539  
 Db 3172 CCAAAATCAGGCTTAAAGACTCCACACTTTTCTTATCAAGGCAATATCCCTTAAGGCC 3231  
 Qy 2540 CAAGGCCCAACAGGAAATCCAAAAGATCTTAAAGGACTTAAAGGCCCAAGGCTTACGTA 2599  
 Db 3232 TGAAGCTCTTAAAGATTTACAGGATATTTGTTAAACATTTAAAGCTCAAGGCTTAGTA 3291  
 Qy 2600 ACCATGCATAACTCCCTGCGAGTAATTCGTTAGTGGATTGAGGAGGACAGAAACCCAGT 2659  
 Db 3292 GAAATGACAGCTCCCTGCAACACCC-----AATCTAGGAGTCAAAAACCAATG 3344  
 Qy 2660 GACAGTGGAGGTTAGTGAAGATCTCAGGATTTAATGAGGAGGCGCTTGTCTTTTATA 2719  
 Db 3345 GTCACTGGA-GACTAGTGAAGATCTTAGACTCATCAAT-GAGGCAAGTAAATCTCTCTATA 3402  
 Qy 2720 CCAGCTGTACCTAGCCCTTACTGTGMYTTCACAAATACACAGAGGAGGAGGAGT 2779  
 Db 3403 TCCAGTTGTACCAACCCCTATACCCCTGCTTTCTCAATATACAGAGGAGGAGGAGT 3462  
 Qy 2780 TACATCTCTGGACCTTTAGGATGCTTCTTGTGCTATCCCTGTACATCTCAGTCTCAAT 2839  
 Db 3463 CATGTTCTGGACCTCAAGGATGCTTCTTCTGTTTCCC-CTGCACTCTGACTCCAGT 3521  
 Qy 2840 CTTGTTGCTTTGAAGATATCTTAAACCCACATCTCAACTCACTGAGCTTTTACC 2899  
 Db 3522 TCTGTTGCTTTGAGGATCCACAGACCACTCCCACTTCAATGATGCTTGTGCTTGC 3581  
 Qy 2900 CCAAGGTTTCAAGGATAGVCCCACTATTTGGCCAGGCAATAGCCCAAGCTTGAGYCA 2959  
 Db 3582 ACAAGGTTTAGGATAGCCCTCAGCTGTTGCTGAGGCACTGGGCCAAGATCTAGGCCA 3641  
 Qy 2960 RTYMTCAATCTGGACACTTCTGCTCTCRGTGAGTGAATTTTACTTTTGGYCCYR 3019  
 Db 3642 CTTCTCAAGTCCAGGCACTTTGGTCTTCAAGTATGTGATGATTTACTTTTGGCTACCA 3701  
 Qy 3020 TTCAGAACTTTGTGCCATCAAGCAACCAAGCCTCTTTMAATTTCTCGYACCTGTGG 3079  
 Db 3702 TTCAGAGGCTCATGCCAGGCTACTCTAGATCTTTGAACCTTTCTACCTAAACAAG 3761  
 Qy 3080 CTACAGGTTTCCAAACSBARGCTCARTCTGTCTCAGAGGAGGTTAATATCTTAGGCT 3139  
 Db 3762 GTACAAGGATCTAGTCAAGGTCAGCTTTGCTTTACAGCAGGCTAAATATCTAGGCT 3821



Db	5964	GTTTCACCAAGGAGATCTAGTGTGGTCAAAATCTCTCCCTCTACCTCCCATCTATGGA	6023
Qy	5354	TACATCCTGGGAAGGACCTTACCCAGTCATTTTATTTATTTACCCCAACTGCGGTTAAAGTGGC	5413
Db	6024	CTCTCTGTGGGAAGGGTCATACTCAATAATCTCTCTACCCACACTGCACTTAAAGGTGGT	6083
Qy	5414	TGGAGTGGAGTCTTTGGATACATCACACTTGGATCAATCTCTGGATCTGCCAAGGAACC	5473
Db	6084	AGGAGTGAATCTTGGATTCAACACCCCGAGTTAAATTTTGGACATCCCTCGAGGAACC	6143
Qy	5474	TGAAATCCAGGAG	5487
Db	6144	TGTGAGACCATCAG	6157

Search completed: August 2, 2006, 02:04:52  
 Job time : 1355 secs

**This Page Blank (uspto)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2006, 17:21:54 ; Search time 6527 Seconds  
(without alignments)  
8099.217 Million cell updates/sec

Title: US-10-717-580-11

Perfect score: 7489.6

Sequence: 1 caacaatgggataataacc.....tattaatacttgcaactgcr 7582

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq 8:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*
- 14: Geneseqn2005s:\*
- 15: Geneseqn2006s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7489.6	100.0	7582	3	AAA59215 Human end
2	7487.4	99.9	7582	2	AAX25665 Complete
3	6448	86.1	56093	6	ABL61744 Colon ade
4	6436	85.9	10499	3	ABN97929 Human ret
5	6124.6	81.8	8523	7	ADS30988 Human gen
6	6124.6	81.8	8523	7	ADP36376 HIRA geno
7	5108	68.2	9502	10	ADF59718 Human con
8	3738.4	49.9	6394	5	AAA59210 DNA encod
9	3431	45.8	4349	5	AAS76475 DNA encod
10	3309.6	44.2	7466	5	AAS68626 DNA encod
11	3049.8	40.7	22436	9	ADA02882 Human PAP
12	3049.8	40.7	22436	10	ADB72620 Human PAP
13	3049.8	40.7	22436	10	ADC85361 Mouse Pap
14	3049.8	40.7	22436	12	ADM74477 Human car
15	3049.8	40.7	22749	14	ADZ12709 Human can
16	2948	39.4	139573	10	ADH58564 Human Na+
17	2948	39.4	156416	13	ADB32817 Human can
18	2932	39.1	5154	5	AAS67609 DNA encod

19	2908.2	38.8	3464	10	ADE09587	Adc09587 Novel DNA
20	2889.6	38.6	2930	6	AAD24195	Aad24195 Human syn
21	2888.6	38.6	2946	2	AAX77526	Aax77526 Human sec
22	2888.6	38.6	2946	3	AAX59468	Aax59468 Human sec
23	2888.6	38.6	2946	10	ADC38776	Adc38776 Human con
24	2734.4	36.5	3831	5	AAS71727	Aas71727 DNA encod
25	2733	36.5	102200	15	AEF74640	Aef74640 Human pol
26	2733	36.5	102200	15	AEF74983	Aef74983 Human pol
27	2688.6	35.9	2938	2	AAX25656	Aax25656 Human end
28	2688.6	35.9	2938	3	AAA59206	Aaa59206 Gag and p
29	2670.8	35.7	173564	13	ABD32953	Abd32953 Human can
30	2613.2	34.9	3372	2	AAX25663	Aax25663 Human end
31	2613.2	34.9	3372	3	AAA59213	Aaa59213 Partial p
32	2550.4	34.1	2599	3	ABN97927	Abn97927 Human ret
33	2541.6	33.9	4535	5	AAS76205	Aas76205 DNA encod
34	2541.6	33.9	4535	10	ADC32196	Adc32196 Human nov
35	2368.4	31.6	8279	5	AAS76474	Aas76474 DNA encod
36	2282.4	30.5	2784	3	ABN97930	Abn97930 Human ret
37	2251.8	30.1	2372	2	AAX25664	Aax25664 Human end
38	2251.8	30.1	2372	3	AAA59214	Aaa59214 R-U5 regi
39	2241	29.9	8294	5	AAS84209	Aas84209 DNA encod
40	2241	29.9	8294	10	ADF59732	Adf59732 Human con
41	2220.4	29.6	161334	11	ACN44334	Acn44334 Human gen
42	2216	29.6	46340	3	ABN97978	Abn97978 Human ret
43	2204.6	29.4	2782	2	AAX25661	Aax25661 Human end
44	2204.6	29.4	2782	3	AAA59211	Aaa59211 5' non co
45	2204.6	29.4	2782	5	AAH20069	Aah20069 HERV-W en

#### ALIGNMENTS

#### RESULT 1

AAA59215

ID AAA59215 standard; DNA; 7582 BP.

XX AC AAA59215;

XX DT 07-NOV-2000 (first entry)

XX DE Human endogenous retrovirus W (HERV-W) sequence.

XX DE Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;

XX KW gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.

XX OS Human endogenous retrovirus.

XX FH Key Location/Qualifiers

FT LTR 1..120

FT /tag= a

FT /note= "R of 5' LTR"

FT LTR 121..575

FT /tag= c

FT /tag= b

FT /note= "U5 of 5' LTR"

FT primer\_bind 579..596

FT CDS 5581..7194

FT /tag= d

FT /note= "ORF1 env538"

FT CDS 7039..7194

FT /tag= e

FT /note= "ORF2 52 AA"

FT CDS 7112..7255

FT /tag= f

FT /note= "ORF3 48 AA"

FT misc\_feature 7244..7254

FT /tag= g

FT /note= "polypurine tract"

FT LTR 7256..7582

FT /tag= h

FT /note= "U3-R of 3' LTR"

FT polyA\_signal 7563..7569

FT /tag= i

```
XX WO200043521-A2.
XX PD 27-JUL-2000.
XX 21-JAN-2000; 2000WO-FR000144.
XX PF 21-JAN-1999; 99FR-00000888.
XX PR 21-JAN-1999; 99FR-00000888.
XX PA (INMR ) BIO MERIEUX.
XX PI Paranhos-Baccala G, Mallet F, Voisset C;
XX DR WPI; 2000-499229/44.
XX PT
XX FT New nucleic acid from human endogenous retrovirus, useful e.g. for
XX PT diagnosis of autoimmune disease and complications of pregnancy, contains
XX PT at least part of the gag gene.
XX PS Disclosure; Page 49-52; 53pp; French.
XX
CC The present sequence represents an endogenous retrovirus, which is
CC associated with an autoimmune disease, and is integrated into the human
CC genome. The retrovirus is human endogenous retrovirus W (HERV-W). The
CC HERV-W retrovirus is associated with autoimmune disease, failure of
CC pregnancy or disorders of pregnancy. HERV-W nucleic acid fragments, or
CC proteins derived from it, are useful for diagnosis of autoimmune disease
CC (specifically multiple sclerosis) and for monitoring pregnancy. The
CC nucleic acid fragments may also be used for in situ labelling of isolated
CC chromosomes, while the transcription product can be used to study or
CC monitor T cell proliferation in vitro
XX
SQ Sequence 7582 BP; 2156 A; 1876 C; 1538 G; 1796 T; 0 U; 216 Other;
Query Match 100.0%; Score 7489.6; DB 3; Length 7582;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAACAATCGGGATATAAACCCAGGCAATTCGAGCTGGCAACAGCAGCCGCCCTTTGGGTCC 60
DB 1 CAACAATCGGGATATAAACCCAGGCAATTCGAGCTGGCAACAGCAGCCGCCCTTTGGGTCC 60
QY 61 CTTCCCTTTGATGGAGCTGTTTCATGCTATTTTCATCTATTAATTTTCACACTGCA 120
DB 61 CTTCCCTTTGATGGAGCTGTTTCATGCTATTTTCATCTATTAATTTTCACACTGCA 120
QY 121 CTTCTCTGCTCCATGTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCCTCCACCACTGC 180
DB 121 CTTCTCTGCTCCATGTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCCTCCACCACTGC 180
QY 181 TGTTTTGCCACCAACCGCANACCTGCGCTGACTTCCCATCCCTCTGGATCCTGCAGGGTGC 240
DB 181 TGTTTTGCCACCAACCGCANACCTGCGCTGACTTCCCATCCCTCTGGATCCTGCAGGGTGC 240
QY 241 CGCTGTGCTCTGATCCAGCGAGCGCCCATTTGCCCTCCCAATTTGGGCTAAAGCTTGC 300
DB 241 CGCTGTGCTCTGATCCAGCGAGCGCCCATTTGCCCTCCCAATTTGGGCTAAAGCTTGC 300
QY 301 CATTGTNCTCGACGGCTAAAGTCCCTGGGTTTGTCTAATTGAGCTGAACACTANTCACT 360
DB 301 CATTGTNCTCGACGGCTAAAGTCCCTGGGTTTGTCTAATTGAGCTGAACACTANTCACT 360
QY 361 GGGTTCATGGTTCCTTTCTGTGACCCAGCGCTTCTAATAKAACTATAACACTTACCACA 420
DB 361 GGGTTCATGGTTCCTTTCTGTGACCCAGCGCTTCTAATAKAACTATAACACTTACCACA 420
QY 421 TGGCCCAAGATTCATTCCTTGGATCCGTGAGGSCAACGAACCTCCAGGTCAGAGATAC 480
DB 421 TGGCCCAAGATTCATTCCTTGGATCCGTGAGGSCAACGAACCTCCAGGTCAGAGATAC 480
QY 481 GARGCTTGGCCACCATCTTGGAGCGGCTGCTACCTCTTGGAGTGGTTTCAACCACATC 540
DB 481 GARGCTTGGCCACCATCTTGGAGCGGCTGCTACCTCTTGGAGTGGTTTCAACCACATC 540
```

```
QY 541 TTGGGAGCTCTGTGAGCAAGGAGCCCGCGTTRACATTTTGGCRACCAAMSRACGACATCC 600
DB 541 TTGGGAGCTCTGTGAGCAAGGAGCCCGCGTTRACATTTTGGCRACCAAMSRACGACATCC 600
QY 601 MAAGTGATGGGAAACGTTTCCCGCAAGACAAAAACGCCCTTAAGACGCTATTCTTGGARAAT 660
DB 601 MAAGTGATGGGAAACGTTTCCCGCAAGACAAAAACGCCCTTAAGACGCTATTCTTGGARAAT 660
QY 661 TGGGAMCAATTTGACCCCTCAGACACATAAGAAAGAAAACGACTTATATTCTTCTGAGATGCC 720
DB 661 TGGGAMCAATTTGACCCCTCAGACACATAAGAAAGAAAACGACTTATATTCTTCTGAGATGCC 720
QY 721 GCCTGGCACTCTGAGGGAAGTATATAATTAACACCATCTTACAGCTAGACACTTCTTTTG 780
DB 721 GCCTGGCACTCTGAGGGAAGTATATAATTAACACCATCTTACAGCTAGACACTTCTTTTG 780
QY 781 TAGAAAAAGGCAAAATGGAGTGAAGTGCATTAAGTACAAAACCTTCTTTTCATTAAGAGACAA 840
DB 781 TAGAAAAAGGCAAAATGGAGTGAAGTGCATTAAGTACAAAACCTTCTTTTCATTAAGAGACAA 840
QY 841 CTCACAAATTATGTAAAAAGTGTGATTTATGCCCTTACAGGAAGCCTTCAGAGTCTACCTCC 900
DB 841 CTCACAAATTATGTAAAAAGTGTGATTTATGCCCTTACAGGAAGCCTTCAGAGTCTACCTCC 900
QY 901 CTATCCCGAGCATCCCGACTCTTCCCGAMYTATTAAGGACCCCGCTTCAACCCCAATGG 960
DB 901 CTATCCCGAGCATCCCGACTCTTCCCGAMYTATTAAGGACCCCGCTTCAACCCCAATGG 960
QY 961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCAAATTTCCCAAT 1020
DB 961 TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCAAATTTCCCAAT 1020
QY 1021 TATGACCCCTCCCAAGCAGTGGGAGGAGAAATTCGGSCCAGCAGAGTGCATGTGCT 1080
DB 1021 TATGACCCCTCCCAAGCAGTGGGAGGAGAAATTCGGSCCAGCAGAGTGCATGTGCT 1080
QY 1081 TTTTCTTCCCGAGCTTAAAGCAATTAACAGACACTTAGTAAATTTCTCAGATAAATCCT 1140
DB 1081 TTTTCTTCCCGAGCTTAAAGCAATTAACAGACACTTAGTAAATTTCTCAGATAAATCCT 1140
QY 1141 GATGGCTATATGTGTGTTTACAGGGTTAGGACAAATTTCTTGATCTGACATGGAGAGAT 1200
DB 1141 GATGGCTATATGTGTGTTTACAGGGTTAGGACAAATTTCTTGATCTGACATGGAGAGAT 1200
QY 1201 ATATATGTCATCTGCTAAATCAGACACTAAACCCCAATGAGAGAAAGTCCACCAATCTGC 1260
DB 1201 ATATATGTCATCTGCTAAATCAGACACTAAACCCCAATGAGAGAAAGTCCACCAATCTGC 1260
QY 1261 AGCCTGAGGTTTGGCGATCTCTGGTATCTCAGTCAAGTCAATGGATANGATGACACA 1320
DB 1261 AGCCTGAGGTTTGGCGATCTCTGGTATCTCAGTCAAGTCAATGGATANGATGACACA 1320
QY 1321 GAAGGAAAGANAATGATTTCCCGCAGCGCAGCAGCTCCAGCTCTASACCCCTCATTTG 1380
DB 1321 GAAGGAAAGANAATGATTTCCCGCAGCGCAGCAGCTCCAGCTCTASACCCCTCATTTG 1380
QY 1381 GGGGACACAGAAATCAGTAACAATGGGAGATTTGGTGTCTGAGACATTTGCTAACTTGTGTGC 1440
DB 1381 GGGGACACAGAAATCAGTAACAATGGGAGATTTGGTGTCTGAGACATTTGCTAACTTGTGTGC 1440
QY 1441 TASAAGGACTAAGGAAACCTASAGAAAGAAATCTATGAATTAATCAATGATGTCACCAATA 1500
DB 1441 TASAAGGACTAAGGAAACCTASAGAAAGAAATCTATGAATTAATCAATGATGTCACCAATA 1500
QY 1501 ACACAGGGGAAGGAAAGAAAATCTACTGCTTTCTGGAGAGACTAAGCGGAGGCAATTCGAG 1560
DB 1501 ACACAGGGGAAGGAAAGAAAATCTACTGCTTTCTGGAGAGACTAAGCGGAGGCAATTCGAG 1560
QY 1561 GAAGCGTGCCTCTCTGTGTCACCTGACTCTTCTGAAAGGCCAACTAAATCTTTAAAGCGTAAGTT 1620
DB 1561 GAAGCGTGCCTCTCTGTGTCACCTGACTCTTCTGAAAGGCCAACTAAATCTTTAAAGCGTAAGTT 1620
```



Qy	1621	TATCACTCAGTCAGCTCAGACATTTAGAAAAAATCTCAAAAGCTCGCGTAGGCCGGAG	1680
Db	1621	TATCACTCAGTCAGCTCAGACATTTAGAAAAAATCTCAAAAGCTCGCGTAGGCCGGAG	1680
Qy	1681	CAAAACTTAGAAACCTTATTGAACTTGGCAACTCGGTCTTTTATATAGAGATCAGGAG	1740
Db	1681	CAAAACTTAGAAACCTTATTGAACTTGGCAACTCGGTCTTTTATATAGAGATCAGGAG	1740
Qy	1741	GAGCAGCGGAACAGGACAAACGGGATTAATAAAGGCCACCGCTTTAGTCATGACCTT	1800
Db	1741	GAGCAGCGGAACAGGACAAACGGGATTAATAAAGGCCACCGCTTTAGTCATGACCTT	1800
Qy	1801	CAGGCAAGTGGACTTTGGAGGCTCTGGAAAGGGAAGCTGGCAAAATTTGAATGCTCTAA	1860
Db	1801	CAGGCAAGTGGACTTTGGAGGCTCTGGAAAGGGAAGCTGGCAAAATTTGAATGCTCTAA	1860
Qy	1861	TAGGGCTTGCTTCCAGTGGGCTCTCAAGGACACTTTAAAAAAGATTTGTCCAAAGTAGAAG	1920
Db	1861	TAGGGCTTGCTTCCAGTGGGCTCTCAAGGACACTTTAAAAAAGATTTGTCCAAAGTAGAAG	1920
Qy	1921	TAAGCGCGCCCTTCGTCCATGCCCTTATTTCAAGGGAATCACTGGAGGCCCACTGCC	1980
Db	1921	TAAGCGCGCCCTTCGTCCATGCCCTTATTTCAAGGGAATCACTGGAGGCCCACTGCC	1980
Qy	1981	CAGGGACAAAGGTCTTTTGGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGACTG	2040
Db	1981	CAGGGACAAAGGTCTTTTGGAGTCAGAAGCCACTAACCAGATGATCCAGCAGCAGACTG	2040
Qy	2041	AGGGTGCTGGGGCAAGCGCATCCCATGCCATCACCTCACAGAGCCCTGGGTATGCTT	2100
Db	2041	AGGGTGCTGGGGCAAGCGCATCCCATGCCATCACCTCACAGAGCCCTGGGTATGCTT	2100
Qy	2101	GACCATGAGGGCCAGGAAGTTGTCTCTGGACACTGGTGGGTCTTCTTAGTCTTACT	2160
Db	2101	GACCATGAGGGCCAGGAAGTTGTCTCTGGACACTGGTGGGTCTTCTTAGTCTTACT	2160
Qy	2161	CTTCTGTCGGGACAACTGTCTCCAGATCTGTCACTATCTCAGGGGTCTCWTAGAGG	2220
Db	2161	CTTCTGTCGGGACAACTGTCTCCAGATCTGTCACTATCTCAGGGGTCTCWTAGAGG	2220
Qy	2221	GGCAGTCACTAGATACCTTCTCCAGCCACTAAGTTATGAATCTGGGGAGCTTTATCTTT	2280
Db	2221	GGCAGTCACTAGATACCTTCTCCAGCCACTAAGTTATGAATCTGGGGAGCTTTATCTTT	2280
Qy	2281	TCACATGCTTTCTAATATGCTTGAAAGCCCACTACCTTGTAGGGAGAGACATTTCTA	2340
Db	2281	TCACATGCTTTCTAATATGCTTGAAAGCCCACTACCTTGTAGGGAGAGACATTTCTA	2340
Qy	2341	GCAAAAGCAGGGGCCATTAACCTGAACATAGGAGAGGAACACCCGTTTGTGTGTC	2400
Db	2341	GCAAAAGCAGGGGCCATTAACCTGAACATAGGAGAGGAACACCCGTTTGTGTGTC	2400
Qy	2401	CCTGCTTGAGGAAGGAATTAATCTCAAGTCTGGGCAACAGAGGACAAATATGGAGG	2460
Db	2401	CCTGCTTGAGGAAGGAATTAATCTCAAGTCTGGGCAACAGAGGACAAATATGGAGG	2460
Qy	2461	CAAGAATCCCGTCTGTTCAAGTTAACTAAGGATTCACCTTCTTTTCCCTACCAAA	2520
Db	2461	CAAGAATCCCGTCTGTTCAAGTTAACTAAGGATTCACCTTCTTTTCCCTACCAAA	2520
Qy	2521	GGCAGTACCCCTCAGACCAAGGCCCAACAGAGGATTCAAAAGATTTGTAAGACTTAA	2580
Db	2521	GGCAGTACCCCTCAGACCAAGGCCCAACAGAGGATTCAAAAGATTTGTAAGACTTAA	2580
Qy	2581	AAGCCCAAGGCTTAGTAAACCATGCACTCCCTGAGTAATTCCTGAGTGGATTGAG	2640
Db	2581	AAGCCCAAGGCTTAGTAAACCATGCACTCCCTGAGTAATTCCTGAGTGGATTGAG	2640
Qy	2641	GAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTTATCAATGG	2700
Db	2641	GAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAAGATCTCAGGATTTATCAATGG	2700
Qy	2701	AGGCGGTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATCTGTGMYTTCCCAATATAC	2760

Db	2701	AGGCGGTGTCCTTTTATACCCAGCTGTACCTAGCCCTTATCTGTGMYTTCCCAATATAC	2760
Qy	2761	CAGAGAAGCAGAGTGTGTTTACASTCCTGTGACCTTMAAGATGCTTCTTCTGATCCCTG	2820
Db	2761	CAGAGAAGCAGAGTGTGTTTACASTCCTGTGACCTTMAAGATGCTTCTTCTGATCCCTG	2820
Qy	2821	TACATCCTGACTCTCAATTTCTGTTTGGCTTTTGAAGATACTTCAAAACCCARCATCTCAAC	2880
Db	2821	TACATCCTGACTCTCAATTTCTGTTTGGCTTTTGAAGATACTTCAAAACCCARCATCTCAAC	2880
Qy	2881	TCACCTGGACTRTTTTACCCCAAGGTTTCAAGGATAGYCCCCATCTATTTTGGCCAGGCAT	2940
Db	2881	TCACCTGGACTRTTTTACCCCAAGGTTTCAAGGATAGYCCCCATCTATTTTGGCCAGGCAT	2940
Qy	2941	TAGCCCAAGACTTGAGYCAATYMTCTACTCTGACACTCTTGTCTTCTTCTGTAAGTGGATG	3000
Db	2941	TAGCCCAAGACTTGAGYCAATYMTCTACTCTGACACTCTTGTCTTCTTCTGTAAGTGGATG	3000
Qy	3001	ATTTACTTTTTRGCGCYRTTTCAGAAACCTTGTGCCATCAAGCCACCCAGRCRCTCTTMA	3060
Db	3001	ATTTACTTTTTRGCGCYRTTTCAGAAACCTTGTGCCATCAAGCCACCCAGRCRCTCTTMA	3060
Qy	3061	ATTTCTCTCGYACCTGTGGCTACAWGGTTTCCAAACSAARAGCTCARCTCTGCTCACAGC	3120
Db	3061	ATTTCTCTCGYACCTGTGGCTACAWGGTTTCCAAACSAARAGCTCARCTCTGCTCACAGC	3120
Qy	3121	AGGTTAAATACCTTAGGCTTAARATTTATCCAAAGGACCCARGGCCCTCAGTGAGGAAYRA	3180
Db	3121	AGGTTAAATACCTTAGGCTTAARATTTATCCAAAGGACCCARGGCCCTCAGTGAGGAAYRA	3180
Qy	3181	TCCAGCTTACTTGGCTTATCTCTCATCYCAAAACCTTAAAGCAACTAAGRRRTTCTTCTG	3240
Db	3181	TCCAGCTTACTTGGCTTATCTCTCATCYCAAAACCTTAAAGCAACTAAGRRRTTCTTCTG	3240
Qy	3241	GRTTAAAGGTTCTGCCGAATGGAATTTCCCAAGTWTGGCRAAATAGCCAGGYCATTA	3300
Db	3241	GRTTAAAGGTTCTGCCGAATGGAATTTCCCAAGTWTGGCRAAATAGCCAGGYCATTA	3300
Qy	3301	WATACASTAATTAAGGAACTCAGAAAGCCCAATACCCATTARTAGATGGAYAMCTGAA	3360
Db	3301	WATACASTAATTAAGGAACTCAGAAAGCCCAATACCCATTARTAGATGGAYAMCTGAA	3360
Qy	3361	GYMAAGTGGCTTCCAGGCCCTTAAAGAGGCCCTTAAACCCAGYCCCAGTGTTAAGYT	3420
Db	3361	GYMAAGTGGCTTCCAGGCCCTTAAAGAGGCCCTTAAACCCAGYCCCAGTGTTAAGYT	3420
Qy	3421	TGCCAACRGGGCAAGACTTTTSTTVATAYRTCAAGAAAAAACAGRAAYAGCTCTRGA	3480
Db	3421	TGCCAACRGGGCAAGACTTTTSTTVATAYRTCAAGAAAAAACAGRAAYAGCTCTRGA	3480
Qy	3481	GTCCCTTACACAGRTCCRAGGAYGAGCTTGAACCCYRTGGCRYACCTGASTAGGAAAYT	3540
Db	3481	GTCCCTTACACAGRTCCRAGGAYGAGCTTGAACCCYRTGGCRYACCTGASTAGGAAAYT	3540
Qy	3541	GATGTAGTGGCAAGGGTTGRCYTCAITTTTAYGGTAGTGTGGCTGCGCAGTACAGTYKTA	3600
Db	3541	GATGTAGTGGCAAGGGTTGRCYTCAITTTTAYGGTAGTGTGGCTGCGCAGTACAGTYKTA	3600
Qy	3601	GTATCTGAAGCAGTTTAAATATAACAGGGAGAGATCTTACTGTGTGGACATCTCATGAK	3660
Db	3601	GTATCTGAAGCAGTTTAAATATAACAGGGAGAGATCTTACTGTGTGGACATCTCATGAK	3660
Qy	3661	GTGAAYRGACATCTCCTCTTAAAGGAGACTTGTGGCTGTGACAAACCTGTTTACTTAAA	3720
Db	3661	GTGAAYRGACATCTCCTCTTAAAGGAGACTTGTGGCTGTGACAAACCTGTTTACTTAAA	3720
Qy	3721	TRTCAAGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGACCTTGTGCAACTCTTAAC	3780
Db	3721	TRTCAAGGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGACCTTGTGCAACTCTTAAC	3780
Qy	3781	CCAGYCNCATTTCTTCCAGACAAATGAAGAAAGATARAAYATACTGTCAACAARTTAAT	3840
Db	3781	CCAGYCNCATTTCTTCCAGACAAATGAAGAAAGATARAAYATACTGTCAACAARTTAAT	3840

3781 CCAGYCNCATTTCTTCCAGACAAATGAGAAAGATARAAYATAAATCTGTCAACAAARTTAATT 3840 Db  
3841 TCTCAAACTATGCCACTCGAGGAGACCTTTTAGAGTTTCTTTGACATGATCCYGCCTT 3900 Qy  
3841 TCTCAAACTATGCCACTCGAGGAGACCTTTTAGAGTTTCTTTGACATGATCCYGCCTT 3900 Db  
3901 CAACCTTGATATCTGATGGAAGTTCTTTGTAGAAAAGGACCTTCGAAAAGYGGGTATGC 3960 Qy  
3901 CAACCTTGATATCTGATGGAAGTTCTTTGTAGAAAAGGACCTTCGAAAAGYGGGTATGC 3960 Db  
3961 AGTGTCTAGTGATATAATGGAATAYTTGAAAGTAATCCCTCTCATCTCAGGAACTAGTGTCTYA 4020 Qy  
3961 AGTGTCTAGTGATATAATGGAATAYTTGAAAGTAATCCCTCTCATCTCAGGAACTAGTGTCTYA 4020 Db  
4021 GCTRCAGAACTAATAGCCYTCAYTKGGGCACTAGAAATAGGAGAAAGRAAAAAGGGYAAA 4080 Qy  
4021 GCTRCAGAACTAATAGCCYTCAYTKGGGCACTAGAAATAGGAGAAAGRAAAAAGGGYAAA 4080 Db  
4081 TATATATACAGACTCTTRATATGCTYACCTAGTCTNTCCATGCCATGMRGCAATATGSAR 4140 Qy  
4081 TATATATACAGACTCTTRATATGCTYACCTAGTCTNTCCATGCCATGMRGCAATATGSAR 4140 Db  
4141 AGAAAAGGAAATTCCTAACTTCYAGRGAAACCTTATCAMACATCAGGAAGCCATTAGGAR 4200 Qy  
4141 AGAAAAGGAAATTCCTAACTTCYAGRGAAACCTTATCAMACATCAGGAAGCCATTAGGAR 4200 Db  
4201 ATTATATYTGCGWGTACAGAAACCTTARAGAGTGMAGTCTTACACTGCGYGGGTATCA 4260 Qy  
4201 ATTATATYTGCGWGTACAGAAACCTTARAGAGTGMAGTCTTACACTGCGYGGGTATCA 4260 Db  
4261 NAAAGGAAGRAAAGGGAATAASAGRGAAATGCTTAAACTTCCCTTAGTATAGGGTAATCC 4320 Qy  
4261 NAAAGGAAGRAAAGGGAATAASAGRGAAATGCTTAAACTTCCCTTAGTATAGGGTAATCC 4320 Db  
4321 TCCAAAGGAGGACCTTCCATTAAGAAATGCTTAAACTTCCCTTAGTATAGGGTAATCC 4380 Qy  
4321 TCCAAAGGAGGACCTTCCATTAAGAAATGCTTAAACTTCCCTTAGTATAGGGTAATCC 4380 Db  
4381 CTTCCGGGAAACCAAGCCCACTACTCAGCAGGAGAAACAGAAATGGGAAACCTTCA 4440 Qy  
4381 CTTCCGGGAAACCAAGCCCACTACTCAGCAGGAGAAACAGAAATGGGAAACCTTCA 4440 Db  
4441 CAGTTTTCTCCCTCGGAGCGGTAGCCACTCAAGAGGAAATACCTTTTCCCTGCAAC 4500 Qy  
4441 CAGTTTTCTCCCTCGGAGCGGTAGCCACTCAAGAGGAAATACCTTTTCCCTGCAAC 4500 Db  
4501 TATCCAAATGGAATTAATTAATTTAATCTTCAATCAACCTTTCACTTAGGCATCGATAGCAC 4560 Qy  
4501 TATCCAAATGGAATTAATTAATTTAATCTTCAATCAACCTTTCACTTAGGCATCGATAGCAC 4560 Db  
4561 CCATCARATGGCCAAATCATTAATTTAATCTGGAACAGGCTTTTCAAACTATCAAGCARAT 4620 Qy  
4561 CCATCARATGGCCAAATCATTAATTTAATCTGGAACAGGCTTTTCAAACTATCAAGCARAT 4620 Db  
4621 AKTCAGGCGCTGTAAGTGTGCAARAAATTAATCCCTGCTCTATCGCCAGAGCTCTTC 4680 Qy  
4621 AKTCAGGCGCTGTAAGTGTGCAARAAATTAATCCCTGCTCTATCGCCAGAGCTCTTC 4680 Db  
4681 AGGAARAAACAAARAAACAGGCCATTAATCTGARAARACTGGCAACTGATTTTACCCACAAG 4740 Qy  
4681 AGGAARAAACAAARAAACAGGCCATTAATCTGARAARACTGGCAACTGATTTTACCCACAAG 4740 Db  
4741 CCCAAACCTCAGGGAATTTTCACTATCTAGTCTGGGTARATACCTTTTCACGGGTGGGA 4800 Qy  
4741 CCCAAACCTCAGGGAATTTTCACTATCTAGTCTGGGTARATACCTTTTCACGGGTGGGA 4800 Db  
4801 RAGGCTTCCCTGTAGGACAGAAAGCCCAAGAGGTAAATAAGGCACTAGTTTCAATGAA 4860 Qy  
4801 RAGGCTTCCCTGTAGGACAGAAAGCCCAAGAGGTAAATAAGGCACTAGTTTCAATGAA 4860 Db  
4861 ATAAATTTCCAGATTCGGAATTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAG 4920 Qy  
4861 ATAAATTTCCAGATTCGGAATTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAG 4920 Db

4921 GCCACAGTAAACCCAGGAGTATCCAGGCGTTAGTATACGATATCACTTACACTGCGCC 4980 Qy  
4921 GCCACAGTAAACCCAGGAGTATCCAGGCGTTAGTATACGATATCACTTACACTGCGCC 4980 Db  
4981 TGAAGGCCACAGTCTCTCAGGGAAGGTGAGAAAAATGAATGAAYACTCAAAAGGACATCTA 5040 Qy  
4981 TGAAGGCCACAGTCTCTCAGGGAAGGTGAGAAAAATGAATGAAYACTCAAAAGGACATCTA 5040 Db  
5041 AAAAGGAAACCCAGGAAACCCACCTCACATGGGCTGTTCTGTGGCTTATAGCCTTAAAA 5100 Qy  
5041 AAAAGGAAACCCAGGAAACCCACCTCACATGGGCTGTTCTGTGGCTTATAGCCTTAAAA 5100 Db  
5101 AGAATCTGCAACTTTCCCAAAAAAGCAGGACTTAGCCCATACGAAATCTGTATGGAAGG 5160 Qy  
5101 AGAATCTGCAACTTTCCCAAAAAAGCAGGACTTAGCCCATACGAAATCTGTATGGAAGG 5160 Db  
5161 CCTTCTATAACCAATGACTTGTGCTTGACCCAAAGACAGCCAACTTAGTTCGAGACATCA 5220 Qy  
5161 CCTTCTATAACCAATGACTTGTGCTTGACCCAAAGACAGCCAACTTAGTTCGAGACATCA 5220 Db  
5221 CCTCTTAGCCAAATATCAACAAGTTCTTAAAAACATTTACAAGGAACCTTATCCCTGAGAAG 5280 Qy  
5221 CCTCTTAGCCAAATATCAACAAGTTCTTAAAAACATTTACAAGGAACCTTATCCCTGAGAAG 5280 Db  
5281 AGGGAAGAAGAACTATTCACCCWMTGACATGGTATTTAGTCAAGTCCCTTCVCTCTAATT 5340 Qy  
5281 AGGGAAGAAGAACTATTCACCCWMTGACATGGTATTTAGTCAAGTCCCTTCVCTCTAATT 5340 Db  
5341 CCCCATCTCTAGATACATCTCTGGGAAGACCTTACCCAGTCAATTTATTTACCCCACTG 5400 Qy  
5341 CCCCATCTCTAGATACATCTCTGGGAAGACCTTACCCAGTCAATTTATTTACCCCACTG 5400 Db  
5401 CGGTTAAAGTGGCTGGAGTGGAGTCTTGGAATACATCACTTTGAGTCAAAATCTCGATAC 5460 Qy  
5401 CGGTTAAAGTGGCTGGAGTGGAGTCTTGGAATACATCACTTTGAGTCAAAATCTCGATAC 5460 Db  
5461 TGCCAAAGAACCTTGAAATCCAGGAGAACAGCTAGCTATTTCTGTGAACTCTTAGAGG 5520 Qy  
5461 TGCCAAAGAACCTTGAAATCCAGGAGAACAGCTAGCTATTTCTGTGAACTCTTAGAGG 5520 Db  
5521 ATTTGGCGCTCTCTTCAAAACAAACCCAGGAGGAAGTAACTAAAAATCAATAATCCCC 5580 Qy  
5521 ATTTGGCGCTCTCTTCAAAACAAACCCAGGAGGAAGTAACTAAAAATCAATAATCCCC 5580 Db  
5581 ATGSGCTCCCTTATCATATTTTCTCTKTASTGTTTTCCTTTTACCTCTCTCACT 5640 Qy  
5581 ATGSGCTCCCTTATCATATTTTCTCTKTASTGTTTTCCTTTTACCTCTCTCACT 5640 Db  
5641 GCACCCCTCCATGCGCTGTATGACCAAGTAGCTCCCTYACCMAGAGTTTCTATGGAGA 5700 Qy  
5641 GCACCCCTCCATGCGCTGTATGACCAAGTAGCTCCCTYACCMAGAGTTTCTATGGAGA 5700 Db  
5701 ATGACGCTCCCGGAAATATTTGATGCCCCATCTGATAGGAGTCTTTTAAAGGAAACCCC 5760 Qy  
5701 ATGACGCTCCCGGAAATATTTGATGCCCCATCTGATAGGAGTCTTTTAAAGGAAACCCC 5760 Db  
5761 ACCTTCACTGCGCCACACCCATATGCCCCGCAACTCTCTACTCTGCGCACTCTTTGCAATG 5820 Qy  
5761 ACCTTCACTGCGCCACACCCATATGCCCCGCAACTCTCTACTCTGCGCACTCTTTGCAATG 5820 Db  
5821 CATGCAAAATCTCAATTTTGGACAGGAAATGATTAATCTCTAGTCTCTGAGGACTT 5880 Qy  
5821 CATGCAAAATCTCAATTTTGGACAGGAAATGATTAATCTCTAGTCTCTGAGGACTT 5880 Db  
5881 GGAGTCACTGTCTGTGGACTTACTTCAACCAAACTGGTATGCTCTGATGGGGTGGAGTT 5940 Qy  
5881 GGAGTCACTGTCTGTGGACTTACTTCAACCAAACTGGTATGCTCTGATGGGGTGGAGTT 5940 Db  
5941 CAAGATCAGGCAAGAGAAAAACATGTAAGAAAGTAATCTCCAACTCAACCGGGGTACAT 6000 Qy  
5941 CAAGATCAGGCAAGAGAAAAACATGTAAGAAAGTAATCTCCAACTCAACCGGGGTACAT 6000 Db

```
QY 6001 GGCACTCTAGCCCTCAAGGACTAGATCTCTCAAACTACATGAAGCCCTCCGTACC 6060
Db 6001 GGCACTCTAGCCCTCAAGGACTAGATCTCTCAAACTACATGAAGCCCTCCGTACC 6060
QY 6061 CATACTCGCTGGTGAAGCCTATTAAATACCAACCTCTAGGGCTCCATGAGGTCTCGGCC 6120
Db 6061 CATACTCGCTGGTGAAGCCTATTAAATACCAACCTCTAGGGCTCCATGAGGTCTCGGCC 6120
QY 6121 CAAAACCTTACTAATCTTGGATATGCTCCCTCGAATCTTCAAGCCATATGTTCAATC 6180
Db 6121 CAAAACCTTACTAATCTTGGATATGCTCCCTCGAATCTTCAAGCCATATGTTCAATC 6180
QY 6181 CTGTGACTCAACCAATGGAACTCTCAGCACAGAAATAAACCACTTCCGTTTGTAGTA 6240
Db 6181 CTGTGACTCAACCAATGGAACTCTCAGCACAGAAATAAACCACTTCCGTTTGTAGTA 6240
QY 6241 GGAACCTCTGTTTCCAAATSGAAATTAACCCATACCTCAAACTCTGTTGAAATTT 6300
Db 6241 GGAACCTCTGTTTCCAAATSGAAATTAACCCATACCTCAAACTCTGTTGAAATTT 6300
QY 6301 AGCAATACTACATACACCAACTCCCAATGATCGGTGGGTAACTCTCTCCCAACAA 6360
Db 6301 AGCAATACTACATACACCAACTCCCAATGATCGGTGGGTAACTCTCTCCCAACAA 6360
QY 6361 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTGCTACCTCAGCCTATCGTCTTTG 6420
Db 6361 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTGCTACCTCAGCCTATCGTCTTTG 6420
QY 6421 AATGGCTCTTCAAGATCTATGCTCTCTCTCAATCTTGTAGTCCCTCCATGRCATCTAC 6480
Db 6421 AATGGCTCTTCAAGATCTATGCTCTCTCTCAATCTTGTAGTCCCTCCATGRCATCTAC 6480
QY 6481 ACTGAACAGAGTTTATACAGTTATGTATCATATCTAAGCCCGCAACAAAGAGTACCATT 6540
Db 6481 ACTGAACAGAGTTTATACAGTTATGTATCATATCTAAGCCCGCAACAAAGAGTACCATT 6540
QY 6541 CTTCTCTTTGTATAGGACAGAGTCTAGTGGCTAGTCTAGTCTGGCATTTGGCGTATC 6600
Db 6541 CTTCTCTTTGTATAGGACAGAGTCTAGTGGCTAGTCTAGTCTGGCATTTGGCGTATC 6600
QY 6601 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAATGGGACATGGAACG 6660
Db 6601 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAATGGGACATGGAACG 6660
QY 6661 GTCCGCGACTCCTGTGTCACTTGAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
Db 6661 GTCCGCGACTCCTGTGTCACTTGAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 6720
QY 6721 CBAATCTGAGAGCTTTAGACTYGTAAACGCTGAPAGAGGGGAACTGTTTATTTTA 6780
Db 6721 CBAATCTGAGAGCTTTAGACTYGTAAACGCTGAPAGAGGGGAACTGTTTATTTTA 6780
QY 6781 GGGGAAGATGCTGTTATTATGTTAATCAATCGGAATCGTCACTGAGAAAGTTAAGAA 6840
Db 6781 GGGGAAGATGCTGTTATTATGTTAATCAATCGGAATCGTCACTGAGAAAGTTAAGAA 6840
QY 6841 ATTCAGATCGAATAACAAGTAKAGCAGARGAGCTTCGAAACACTGGACCTGGGCGCTC 6900
Db 6841 ATTCAGATCGAATAACAAGTAKAGCAGARGAGCTTCGAAACACTGGACCTGGGCGCTC 6900
QY 6901 CTCAGCCATGAGTCCCTGGATCTCCCTCTTTAGGACCTCTAGCAGCTATATATTG 6960
Db 6901 CTCAGCCATGAGTCCCTGGATCTCCCTCTTTAGGACCTCTAGCAGCTATATATTG 6960
QY 6961 CTACTCTCTTTGGACCTGTATCTTTTACCTCTCTGTTAACTTTGTCTCTCCGAAATC 7020
Db 6961 CTACTCTCTTTGGACCTGTATCTTTTACCTCTCTGTTAACTTTGTCTCTCCGAAATC 7020
QY 7021 GAAGCTGTAAACTACAAATGGAGCCCAAGATGCAAGTCAAGATCTACCGCAGA 7080
Db 7021 GAAGCTGTAAACTACAAATGGAGCCCAAGATGCAAGTCAAGATCTACCGCAGA 7080
QY 7081 CCCCTGGACCGGCTGYTAGCCCAAGATCTGATGTTAATGACATCAAGGCAACCCCTCT 7140
```

```
Db 7081 CCCCTGGACCGGCTGYTAGCCCAAGATCTGATGTTAATGACATCAAGGCAACCCCTCT 7140
QY 7141 GAGGAATCTCAGCTGCACAACCTCTACTACGCCCAATTTACGAGGAACGATTTAGAGC 7200
Db 7141 GAGGAATCTCAGCTGCACAACCTCTACTACGCCCAATTTACGAGGAACGATTTAGAGC 7200
QY 7201 GGTGTGGCCAACTCTCCCAACAGCACTTAGGTCTTCTGTGTTGAGATGGGGACTGAGA 7260
Db 7201 GGTGTGGCCAACTCTCCCAACAGCACTTAGGTCTTCTGTGTTGAGATGGGGACTGAGA 7260
QY 7261 GACAGACTAGCTGGATTTCTTAGGCTGATTAAGAATCCYTAAGCCTAGSTGGGAAGGTG 7320
Db 7261 GACAGACTAGCTGGATTTCTTAGGCTGATTAAGAATCCYTAAGCCTAGSTGGGAAGGTG 7320
QY 7321 ACCACATCCACCTTTAAACAACGGGGCTTGCAACTTAGYTACACCTGACCAATCAATTCG 7380
Db 7321 ACCACATCCACCTTTAAACAACGGGGCTTGCAACTTAGYTACACCTGACCAATCAATTCG 7380
QY 7381 CTCACATAAATGCTAATAGGCAAGACAGAGGTAAAGAAATAGCAATCATYTTATTC 7440
Db 7381 CTCACATAAATGCTAATAGGCAAGACAGAGGTAAAGAAATAGCAATCATYTTATTC 7440
QY 7441 MTGAGAGCACAGAGGAGGCAATGATCGGATATAAAACCCAAAGTTCGAGCCGGCAA 7500
Db 7441 MTGAGAGCACAGAGGAGGCAATGATCGGATATAAAACCCAAAGTTCGAGCCGGCAA 7500
QY 7501 CGGCAACCCCTTTGGGTCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTTCAC 7560
Db 7501 CGGCAACCCCTTTGGGTCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTTCAC 7560
QY 7561 TCTATTAAATCTTGCACTGCR 7582
Db 7561 TCTATTAAATCTTGCACTGCR 7582

RESULT 2
AX25665
ID AX25665 standard; cDNA to mRNA; 7582 BP.
XX AC AX25665;
XX DT 21-MAY-1999 (first entry)
XX Complete human endogenous retrovirus W genome.
XX DE
XX KW Clone; human endogenous retrovirus; genome; autoimmune disease;
XX KW multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
XX KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX OS Human endogenous retrovirus.
XX PN WO9902696-A1.
XX PD 21-JAN-1999.
XX PF 06-JUL-1998; 98WO-FR001442.
XX PR 07-JUL-1997; 97FR-00008815.
XX PA (INMR ) BIO MERIEUX.
XX PI Beseme F, Blond J, Bouton O, Mandrand B, Mallet F;
XX DR WPI; 1999-120897/10.
XX PT New nucleic acid sequences from human endogenous retrovirus-W - expressed
XX PT exclusively in placenta and useful in diagnosis and therapy of autoimmune
XX PS disease, and abnormal or failed pregnancy.
XX Claim 1; Page 71-74; 106pp; French.
XX CC This sequence represents the complete sequence of the human endogenous
```

CC retrovirus (HERV) W genome. The nucleic acids, their fragments or  
 CC peptides encoded by them are markers of autoimmune disease (e.g. multiple  
 CC sclerosis, rheumatoid polyarthritis, disseminated lupus erythematosus,  
 CC insulin-dependent diabetes and related pathologies) and of abnormal or  
 CC unsuccessful pregnancy and can be used as chromosomal markers for  
 CC susceptibility to these conditions, or proximity markers of genes  
 CC associated with this susceptibility  
 XX

SQ Sequence 7582 BP; 2156 A; 1877 C; 1537 G; 1796 T; 2 U; 214 Other;

Query Match 99.9%; Score 7487.4; DB 2; Length 7582;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 7578; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy	1	CAACAATCGGGATATAAACCCAGGACATTCGAGCTGGCAACGACGCCCCCTTTGGGTCC	60
Db	1	CAACAATCGGGATATAAACCCAGGACATTCGAGCTGGCAACGACGCCCCCTTTGGGTCC	60
Qy	61	CTTCCCTTTGATGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCACACTGCA	120
Db	61	CTTCCCTTTGATGGAGCTGTTTTCATGCTATTTCACTCTATTAAATCTTGCACACTGCA	120
Qy	121	CTCTCTCGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC	180
Db	121	CTCTCTCGTCCATGTTTCTTACGGCTCGAGCTGAGCTTTTGCTCACCGTCCACCACTGC	180
Qy	181	TGTTTGGCCACCAACCCANACCTGCGCTGACTCCCATCCCTTGGATCCTGCGAGGTGTC	240
Db	181	TGTTTGGCCACCAACCCANACCTGCGCTGACTCCCATCCCTTGGATCCTGCGAGGTGTC	240
Qy	241	CGCTGTGCTCTGATCCAGCGARGCGCCATTGCGCGCTCCCAATTGGGCTAAAGGCTTGC	300
Db	241	CGCTGTGCTCTGATCCAGCGARGCGCCATTGCGCGCTCCCAATTGGGCTAAAGGCTTGC	300
Qy	301	CATTGTNCCGTGACCGGCTAAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACT	360
Db	301	CATTGTNCCGTGACCGGCTAAAGTGCCTGGGTTTGTTCTAATTGAGCTGAACACTANTCACT	360
Qy	361	GGGTTCCATGGTCTCTTGTGACCCAGCGGCTCTAATKAACCTATAACACTTACCACA	420
Db	361	GGGTTCCATGGTCTCTTGTGACCCAGCGGCTCTAATKAACCTATAACACTTACCACA	420
Qy	421	TGGCCCAAGATTCATCTCTTGGAAATCCGTGAGGCAACGAACCTCCAGTCAAGAAATAC	480
Db	421	TGGCCCAAGATTCATCTCTTGGAAATCCGTGAGGCAACGAACCTCCAGTCAAGAAATAC	480
Qy	481	GARGCTTGCACCATCTTTGGAAAGCGGCTGTACCTCTTGGAAAGTGTTCACCACTC	540
Db	481	GARGCTTGCACCATCTTTGGAAAGCGGCTGTACCTCTTGGAAAGTGTTCACCACTC	540
Qy	541	TTGGGAGCTCTGTGACGAGGACCCCGGTTACATTTTGGCRAACMSRACGGACATCC	600
Db	541	TTGGGAGCTCTGTGACGAGGACCCCGGTTACATTTTGGCRAACMSRACGGACATCC	600
Qy	601	MAAGTGATGGGAAACCTTCCCGCAAGACAAAACGCCCTTAAGACGTATTCTGGAAT	660
Db	601	MAAGTGATGGGAAACCTTCCCGCAAGACAAAACGCCCTTAAGACGTATTCTGGAAT	660
Qy	661	TGGGAMCAATTTGACCTTCAGACACTAAGAAAGAAACGACTTATTTCTTGCAGTGCC	720
Db	661	TGGGAMCAATTTGACCTTCAGACACTAAGAAAGAAACGACTTATTTCTTGCAGTGCC	720
Qy	721	GCTTGCACTCTGTGAGGAGATATAATTAACACCATCTTACAGCTAGACTCTTTTG	780
Db	721	GCTTGCACTCTGTGAGGAGATATAATTAACACCATCTTACAGCTAGACTCTTTTG	780
Qy	781	TAGAAAAGCAATGGAGTGAAGTGCCATAAGTACAAAACCTTTCTTCAATTAAGACAAA	840
Db	781	TAGAAAAGCAATGGAGTGAAGTGCCATAAGTACAAAACCTTTCTTCAATTAAGACAAA	840
Qy	841	CTCACAAATTAATGTAAGAGTGTGATTTATGCCCCCTACAGGAAGCCTTCAGAGTCTACCTCC	900
Db	841	CTCACAAATTAATGTAAGAGTGTGATTTATGCCCCCTACAGGAAGCCTTCAGAGTCTACCTCC	900

Qy	901	CTATCCAGCATCCCGACTCTTCCCACMYTAATAAGGACCCCTTCAACCCAAATGG	960
Db	901	CTATCCAGCATCCCGACTCTTCCCACMYTAATAAGGACCCCTTCAACCCAAATGG	960
Qy	961	TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCAAATATTTCCCAAT	1020
Db	961	TCCAAAAGGAGATAGACAAAAGGGTAAACAGTGAACCAAGAGTGCCAAATATTTCCCAAT	1020
Qy	1021	TATGACCCCTCCCAGCAGTGGGAGGAGAGAAATTCGSCCCAGCAGAGTGCATGTGCVT	1080
Db	1021	TATGACCCCTCCCAGCAGTGGGAGGAGAGAAATTCGSCCCAGCAGAGTGCATGTGCVT	1080
Qy	1081	TTTTTCTTCCAGACTTTAAAGCAATAAAAAACAGACTTAGGTAAATCTCAGATAAAYCCT	1140
Db	1081	TTTTTCTTCCAGACTTTAAAGCAATAAAAAACAGACTTAGGTAAATCTCAGATAAAYCCT	1140
Qy	1141	GATGGCTATATTGRTGTTTACAAGGGTTAGGACAAATCTTTGATCTGACATGGAGAGAT	1200
Db	1141	GATGGCTATATTGRTGTTTACAAGGGTTAGGACAAATCTTTGATCTGACATGGAGAGAT	1200
Qy	1201	ATATATGTCATGCTAAATCAGACACTAAACCCCAATGAGAGAGTGCCACCACTAATGC	1260
Db	1201	ATATATGTCATGCTAAATCAGACACTAAACCCCAATGAGAGAGTGCCACCACTAATGC	1260
Qy	1261	AGCCTGAGRGTTTGGCGATCTCTGATCTCAGTCAGTCAATGATGATGATGATGATGATG	1320
Db	1261	AGCCTGAGRGTTTGGCGATCTCTGATCTCAGTCAGTCAATGATGATGATGATGATGATG	1320
Qy	1321	GAAGGAAAGANAATGATTTCCCAACAGGCGAGCAGTCCAGTCTTASACCTCATG	1380
Db	1321	GAAGGAAAGANAATGATTTCCCAACAGGCGAGCAGTCCAGTCTTASACCTCATG	1380
Qy	1381	GGGACACAGAAATCAGTAACATGGAGATTTGGTCTGCGAGACTTTGCTTAATGCTGTC	1440
Db	1381	GGGACACAGAAATCAGTAACATGGAGATTTGGTCTGCGAGACTTTGCTTAATGCTGTC	1440
Qy	1441	TASAGGACTAAGGAAACTASGAGAAATCTATGAATTAATCAATGATGTCACCATTA	1500
Db	1441	TASAGGACTAAGGAAACTASGAGAAATCTATGAATTAATCAATGATGTCACCATTA	1500
Qy	1501	ACACAGGAAAGGAAAGAAATCTTCTGCTTTTGGAGAGACTAAGGAGGAGATTTGAG	1560
Db	1501	ACACAGGAAAGGAAAGAAATCTTCTGCTTTTGGAGAGACTAAGGAGGAGATTTGAG	1560
Qy	1561	GAAGGCTGCTCTGTGACCTGACTCTTCTGAAGGCCAACTAACTTTAAAGCGTAAGTT	1620
Db	1561	GAAGGCTGCTCTGTGACCTGACTCTTCTGAAGGCCAACTAACTTTAAAGCGTAAGTT	1620
Qy	1621	TATCACTCAGTCAGTCGACACTTTAGAAAAAACTTCAAAAAGTCTGCGTAGGCCCGGAG	1680
Db	1621	TATCACTCAGTCAGTCGACACTTTAGAAAAAACTTCAAAAAGTCTGCGTAGGCCCGGAG	1680
Qy	1681	CAAAAATTAGAAAACCTATTGAACTTGGCAACUTCGGTCTTTTATAATAGAGATCAGGAG	1740
Db	1681	CAAAAATTAGAAAACCTATTGAACTTGGCAACUTCGGTCTTTTATAATAGAGATCAGGAG	1740
Qy	1741	GAGCAGGCGGAACAGGCAAAAGGGATTAAGAAAAGGCCACCGCTTTAGTCAATGACCT	1800
Db	1741	GAGCAGGCGGAACAGGCAAAAGGGATTAAGAAAAGGCCACCGCTTTAGTCAATGACCT	1800
Qy	1801	CAGGCAAGTGCATTTGGAGGCTCTGGAAGGAAAAAGTGGGCAAAATTTGAATGCTTAA	1860
Db	1801	CAGGCAAGTGCATTTGGAGGCTCTGGAAGGAAAAAGTGGGCAAAATTTGAATGCTTAA	1860
Qy	1861	TAGGCTTGTCTTCCAGTCCGCTCTACAAAGGACATTTTAAAAAAGATTTGTCGAAGTAGAG	1920
Db	1861	TAGGCTTGTCTTCCAGTCCGCTCTACAAAGGACATTTTAAAAAAGATTTGTCGAAGTAGAG	1920
Qy	1921	TAAAGCGGCCCTTGTCTCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCACCTGCC	1980
Db	1921	TAAAGCGGCCCTTGTCTCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCACCTGCC	1980

1981 QY CAGGGGACAAAGGTCCTTTGAGTCAGAAAGCCACTTAACCCAGATGATCCAGCAGCAGGACTG 2040  
1981 Db CAGGGGACAAAGGTCCTTTGAGTCAGAAAGCCACTTAACCCAGATGATCCAGCAGCAGGACTG 2040  
2041 QY AGGGTGCTGGGGAGCGGCATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTT 2100  
2041 Db AGGGTGCTGGGGAGCGGCATCCCATGCCATCACCCTCACAGAGCCCTGGGTATGCTTT 2100  
2101 QY GACCATTTGAGGGCCAGGAAGGTTGTCTCTGGACACTGGTGCCTTTCTTAGCTTTACT 2160  
2101 Db GACCATTTGAGGGCCAGGAAGGTTGTCTCTGGACACTGGTGCCTTTCTTAGCTTTACT 2160  
2161 QY CTTCTGTCCCGGACCACTGTCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGAGC 2220  
2161 Db CTTCTGTCCCGGACCACTGTCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGAGC 2220  
2221 QY GGCAGTCACCTAGATACCTTTTCCAGCACCTAAGTTATGAATGGGAGCTTTATCTTTT 2280  
2221 Db GGCAGTCACCTAGATACCTTTTCCAGCACCTAAGTTATGAATGGGAGCTTTATCTTTT 2280  
2281 QY TCACATGCTTTTCTTAATTATGCTTGAAGCCCACTACCTTGTAGGGAGACATTTCTA 2340  
2281 Db TCACATGCTTTTCTTAATTATGCTTGAAGCCCACTACCTTGTAGGGAGACATTTCTA 2340  
2341 QY GCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAGGAACACCCGTTTCTGTNCC 2400  
2341 Db GCAAAAGCAGGGGCCATTATACACCTGAACATAGGAGAGGAACACCCGTTTCTGTNCC 2400  
2401 QY CTTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAGAGCAATATGACGAGC 2460  
2401 Db CTTGCTTGAGGAAGGAATTAATCCTGAAGTCTGGGCAACAGAGAGCAATATGACGAGC 2460  
2461 QY CAAAGAAATGCCCTGCTGTTCAAGTTAAACTAAAGGATCCACTTCTTTTCCCTACCAAA 2520  
2461 Db CAAAGAAATGCCCTGCTGTTCAAGTTAAACTAAAGGATCCACTTCTTTTCCCTACCAAA 2520  
2521 QY GGCAGTACCCCTCAGACCCCAAGGCCCAAGGATTCCAAAGATTTGTTAAGGACTTAA 2580  
2521 Db GGCAGTACCCCTCAGACCCCAAGGCCCAAGGATTCCAAAGATTTGTTAAGGACTTAA 2580  
2581 QY AAGCCCAAGGCTTAGTAAACCACTGATTAACCTCCCTGCACTAATCCGTAGTGGATTGAG 2640  
2581 Db AAGCCCAAGGCTTAGTAAACCACTGATTAACCTCCCTGCACTAATCCGTAGTGGATTGAG 2640  
2641 QY GAGGCA CAGAAACCCAGTGGACAGTGGAGGTTAGTGCAGATCTCAGGATTTCAATGG 2700  
2641 Db GAGGCA CAGAAACCCAGTGGACAGTGGAGGTTAGTGCAGATCTCAGGATTTCAATGG 2700  
2701 QY AGGCCGTTGCTTTTATACCCAGCTGTACCTAGCCCTTATCTGTGMYTTCCCAAAATAC 2760  
2701 Db AGGCCGTTGCTTTTATACCCAGCTGTACCTAGCCCTTATCTGTGMYTTCCCAAAATAC 2760  
2761 QY CAGAGAAGCAGAGTGGTTTACASTCTCGACCTTMAAGATGCCTTCTTCGATCCCTG 2820  
2761 Db CAGAGAAGCAGAGTGGTTTACASTCTCGACCTTMAAGATGCCTTCTTCGATCCCTG 2820  
2821 QY TACATCCTGACTCTCAATTTCTTTGCTTTTGAAGATCTTCAAAACCCARCATCTCAAC 2880  
2821 Db TACATCCTGACTCTCAATTTCTTTGCTTTTGAAGATCTTCAAAACCCARCATCTCAAC 2880  
2881 QY TCACCTGGACTRTTTTACCCCAAGGTTTCAAGGATAGYCCCATCTATTATGGCCAGGCAT 2940  
2881 Db TCACCTGGACTRTTTTACCCCAAGGTTTCAAGGATAGYCCCATCTATTATGGCCAGGCAT 2940  
2941 QY TAGCCCAAGACTTGAGYCACTTATACCTGACACTCTTGTCTTCTGTAGTGGATG 3000  
2941 Db TAGCCCAAGACTTGAGYCACTTATACCTGACACTCTTGTCTTCTGTAGTGGATG 3000  
3001 QY ATTTACTTTTTRGCGCCYRTTCAGAAAACCTTGTGCCATCAAGCCACCCAGGCRCTTTMA 3060  
3001 Db ATTTACTTTTTRGCGCCYRTTCAGAAAACCTTGTGCCATCAAGCCACCCAGGCRCTTTMA 3060  
3061 QY ATTTCTCTCGCYACTGTGGCTACAWGGTTTCCAAAC SARARGCTCARCTCTGCTCACAGC 3120

3061 Db ATTTCTCTCGCYACTGTGGCTACAWGGTTTCCAAAC SARARGCTCARCTCTGCTCACAGC 3120  
3121 QY AGGTTAAATACCTTAGGCTTAARATTTCAAAGGCAACCAAGGCCCTCAGTGAGGAAYRYA 3180  
3121 Db AGGTTAAATACCTTAGGCTTAARATTTCAAAGGCAACCAAGGCCCTCAGTGAGGAAYRYA 3180  
3181 QY TCCAGGCTATACCTAGGCTTTATCTCTCATCYCAAAACCCCTAAAGCAACTAAGRRRTTCTTTG 3240  
3181 Db TCCAGGCTATACCTAGGCTTTATCTCTCATCYCAAAACCCCTAAAGCAACTAAGRRRTTCTTTG 3240  
3241 QY GCRTAAYAGGTTTCTCCGGAATGATTTCCCGAGGTGTGGCRAAATAGCCAGGYCATTA 3300  
3241 Db GCRTAAYAGGTTTCTCCGGAATGATTTCCCGAGGTGTGGCRAAATAGCCAGGYCATTA 3300  
3301 QY WATACASTAATTAAGGAAACTCAGAAAGCAATACCCATTTTARTTAAGATGGAYAMCTGAA 3360  
3301 Db WATACASTAATTAAGGAAACTCAGAAAGCAATACCCATTTTARTTAAGATGGAYAMCTGAA 3360  
3361 QY GYMRAGTGGCTTTCCAGGCCCTTAAAGAGGCTTTAAACCCCAAGYCCCAAGTGTTAAGYT 3420  
3361 Db GYMRAGTGGCTTTCCAGGCCCTTAAAGAGGCTTTAAACCCCAAGYCCCAAGTGTTAAGYT 3420  
3421 QY TGCCAACRGGSCAAGACTTTTSTTTATATYRTACAGAAAAAAGCAGRAAYAGCTCTRGA 3480  
3421 Db TGCCAACRGGSCAAGACTTTTSTTTATATYRTACAGAAAAAAGCAGRAAYAGCTCTRGA 3480  
3481 QY GTCTCTTACACAGTCCRAGGGAAGCTTTGCAACCCVTCGCRVACCTGASTAAGGAAAT 3540  
3481 Db GTCTCTTACACAGTCCRAGGGAAGCTTTGCAACCCVTCGCRVACCTGASTAAGGAAAT 3540  
3541 QY GATGTAGTGCCAAAGGTTGRCYTCATTTGTTTAYGGGTAGTGTGTCAGTACAGTYKTA 3600  
3541 Db GATGTAGTGCCAAAGGTTGRCYTCATTTGTTTAYGGGTAGTGTGTCAGTACAGTYKTA 3600  
3601 QY GTATCTGAAGCAGTTTAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
3601 Db GTATCTGAAGCAGTTTAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAK 3660  
3661 QY GTGAAYRGCATACTCACTCTAAAGGAGACTTGTGGCTGTGACAAACCTTACTTAA 3720  
3661 Db GTGAAYRGCATACTCACTCTAAAGGAGACTTGTGGCTGTGACAAACCTTACTTAA 3720  
3721 QY TRTCAAGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACTTGTGCAACTCTTAAC 3780  
3721 Db TRTCAAGCTCTATTACTTGAARGGCCAGTGTGCRACCTGTGCACTTGTGCAACTCTTAAC 3780  
3781 QY CCAGYCNCAATTTCTTCCAGCAATGAAGAAAAAGATARAAYATACTGTCAACAARTAAAT 3840  
3781 Db CCAGYCNCAATTTCTTCCAGCAATGAAGAAAAAGATARAAYATACTGTCAACAARTAAAT 3840  
3841 QY TCTCAAAACCTATGCCACTCGAGGGGACCTTTTAGAGTTCCYTTGACTGATCCYGNACCTT 3900  
3841 Db TCTCAAAACCTATGCCACTCGAGGGGACCTTTTAGAGTTCCYTTGACTGATCCYGNACCTT 3900  
3901 QY CAACTTGTATACTGATGGAAGTTCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGC 3960  
3901 Db CAACTTGTATACTGATGGAAGTTCTTTGTAGAAAAAGGACTTCGAAAAGYGGGGTATGC 3960  
3961 QY AGTGTGCTAGTGAATAATGAATAYTTTGAAGTAATCCCTCACTCCAGGAACCTAGTCTYA 4020  
3961 Db AGTGTGCTAGTGAATAATGAATAYTTTGAAGTAATCCCTCACTCCAGGAACCTAGTCTYA 4020  
4021 QY GCTTRCAGAACTAATAGCCYTCAYTGGGCACTAGAAATAGGAGAAAGAAAAAGGGYAAA 4080  
4021 Db GCTTRCAGAACTAATAGCCYTCAYTGGGCACTAGAAATAGGAGAAAGAAAAAGGGYAAA 4080  
4081 QY TATATATACAGCTCTTRATATGCTYACCTAGTCTCATGCTCCCATMRGCAATATGSAR 4140  
4081 Db TATATATACAGCTCTTRATATGCTYACCTAGTCTCATGCTCCCATMRGCAATATGSAR 4140  
4141 QY AGAAAGGGAATTCCTAACTTCYAGRGGAACCTTATCAMAATCAGGAAGCCATTAGGAR 4200



Db 4141 AGAAGGAAATTCCTAACTTCYAGRGAAACACCTATCAMACATCAGGAAGCCATTAGGAR 4200  
Qy 4201 ATTATTATYGGWGTACAGAAACCTARAGAGTGGWAGTCTTACACTGCGYGGGTCACTCA 4260  
Db 4201 ATTATTATYGGWGTACAGAAACCTARAGAGTGGWAGTCTTACACTGCGYGGGTCACTCA 4260  
Qy 4261 NAAAGGAAAGRAAGGGAAATASAGRGAAATGCGCAAGCAKATATGGAAGWMAAAGAGC 4320  
Db 4261 NAAAGGAAAGRAAGGGAAATASAGRGAAATGCGCAAGCAKATATGGAAGWMAAAGAGC 4320  
Qy 4321 TGCAAGCAGGACCCCTCCATTAGAAATGCTTATTAACCTTCCCTTAGTATAGGGTAATCC 4380  
Db 4321 TGCAAGCAGGACCCCTCCATTAGAAATGCTTATTAACCTTCCCTTAGTATAGGGTAATCC 4380  
Qy 4381 CTTCCGGGAAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCAGGAGG 4440  
Db 4381 CTTCCGGGAAACCAAGCCCGAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCAGGAGG 4440  
Qy 4441 CAGTTTTCTCCCTCGGGACGGTTAGCCACTGGAAGGGAATACTTTTTGCCCTGCAAC 4500  
Db 4441 CAGTTTTCTCCCTCGGGACGGTTAGCCACTGGAAGGGAATACTTTTTGCCCTGCAAC 4500  
Qy 4501 TATCAATGGAAATTAATTAACACCTTTCATCAAACTTTTCACTTTAGGCATCGATAGCAC 4560  
Db 4501 TATCAATGGAAATTAATTAACACCTTTCATCAAACTTTTCACTTTAGGCATCGATAGCAC 4560  
Qy 4561 CCATCARATGGCAATCATATTTACTGGACAGGCCCTTTTCAAACTATCAAGCARAT 4620  
Db 4561 CCATCARATGGCAATCATATTTACTGGACAGGCCCTTTTCAAACTATCAAGCARAT 4620  
Qy 4621 AKTCAGGGCTGTGAATGTGCAARAATAATCCCTGCCCTYATGCGCAAGCTCCCTTC 4680  
Db 4621 AKTCAGGGCTGTGAATGTGCAARAATAATCCCTGCCCTYATGCGCAAGCTCCCTTC 4680  
Qy 4681 AGGAAACAAARAAACAGGCCATTACCTGARAARACTGGCAACTGATTTTACCCACAAG 4740  
Db 4681 AGGAAACAAARAAACAGGCCATTACCTGARAARACTGGCAACTGATTTTACCCACAAG 4740  
Qy 4741 CCCAAACCTCAGGGATTTTCAGTATCTA CTAGTCTGGGTARATAC TTTTCA CGGGTTGGCA 4800  
Db 4741 CCCAAACCTCAGGGATTTTCAGTATCTA CTAGTCTGGGTARATAC TTTTCA CGGGTTGGCA 4800  
Qy 4801 RAGGCTTCCCTGTAGGACAGAAAGGCCCAAGAGTAAATAAGGCACTAGTTCATGAA 4860  
Db 4801 RAGGCTTCCCTGTAGGACAGAAAGGCCCAAGAGTAAATAAGGCACTAGTTCATGAA 4860  
Qy 4861 ATAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAG 4920  
Db 4861 ATAATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAG 4920  
Qy 4921 GCCACAGTAAACCAAGGAGTATCCCAAGCGTTAGGTATACGATATCACTTACCTGGGCC 4980  
Db 4921 GCCACAGTAAACCAAGGAGTATCCCAAGCGTTAGGTATACGATATCACTTACCTGGGCC 4980  
Qy 4981 TGAAGGCCACAGTCTCAGGAGGTTCGAGAAATGAATGAAYACTCAAGGACATCTA 5040  
Db 4981 TGAAGGCCACAGTCTCAGGAGGTTCGAGAAATGAATGAAYACTCAAGGACATCTA 5040  
Qy 5041 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCCCTGYTCTGTGTGCTATAGCCCTTAAA 5100  
Db 5041 AAAAGCAAAACCCAGGAAACCCACCTCACATGGCCCTGYTCTGTGTGCTATAGCCCTTAAA 5100  
Qy 5101 AGAATCTGCAATTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGG 5160  
Db 5101 AGAATCTGCAATTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGG 5160  
Qy 5161 CCTTTCATACCAATGACCTTGCTTGACCCCAAGCAGCCACTTAGTTGACAGATCA 5220  
Db 5161 CCTTTCATACCAATGACCTTGCTTGACCCCAAGCAGCCACTTAGTTGACAGATCA 5220  
Qy 5221 CCTCTTAGCCAAATATCAACAGTTCTTAAAAATTACAAGGAACCTATCCCTGAGAAG 5280  
Db 5221 CCTCTTAGCCAAATATCAACAGTTCTTAAAAATTACAAGGAACCTATCCCTGAGAAG 5280

Qy 5281 AGGAAAAGAACTATTCCACCCWMTGACATGGTATTAGTCAAGTCCCTTCVCTCTAATT 5340  
Db 5281 AGGAAAAGAACTATTCCACCCWMTGACATGGTATTAGTCAAGTCCCTTCVCTCTAATT 5340  
Qy 5341 CCCCATCCCTAGATACATCTCTGGGAAGACCCCTACCCAGTCAATTTATTTATACCCCACTG 5400  
Db 5341 CCCCATCCCTAGATACATCTCTGGGAAGACCCCTACCCAGTCAATTTATTTATACCCCACTG 5400  
Qy 5401 CGGTTAAAGTGGCTGGAGTGGAGTCTTGATACATCACTTTGAGTCAAAATCCTGGATAC 5460  
Db 5401 CGGTTAAAGTGGCTGGAGTGGAGTCTTGATACATCACTTTGAGTCAAAATCCTGGATAC 5460  
Qy 5461 TGCCAAAGGAACCTGAAATCCAGGAGACACGCTAGCTATTCCTGTGCAACCTCTAGAGG 5520  
Db 5461 TGCCAAAGGAACCTGAAATCCAGGAGACACGCTAGCTATTCCTGTGCAACCTCTAGAGG 5520  
Qy 5521 ATTTGGCGCTCTCTTCAAAACAAACACAGGAGGAAAGTAATAAAATCATAAATCCCCC 5580  
Db 5521 ATTTGGCGCTCTCTTCAAAACAAACACAGGAGGAAAGTAATAAAATCATAAATCCCCC 5580  
Qy 5581 ATGSGCCTCCTTATCATATTTTCTCTKTASTGTSTTTTACCCTSTTTCACTCTCACT 5640  
Db 5581 ATGSGCCTCCTTATCATATTTTCTCTKTASTGTSTTTTACCCTSTTTCACTCTCACT 5640  
Qy 5641 GCACCCCTCCATGCCCTGTATGACACGATAGCTCCCTYACCMAGAGTTTCTATGGAGA 5700  
Db 5641 GCACCCCTCCATGCCCTGTATGACACGATAGCTCCCTYACCMAGAGTTTCTATGGAGA 5700  
Qy 5701 ATGCAAGCTCCCGGAAATATTTGATGCCCCATCTGATAGGAGTCTTTTAAAGGGAACCCC 5760  
Db 5701 ATGCAAGCTCCCGGAAATATTTGATGCCCCATCTGATAGGAGTCTTTTAAAGGGAACCCC 5760  
Qy 5761 ACCTTCACTGCCCAACCCATATGATGCCCACTCGTATAGGAGTCTTTTAAAGGGAACCCC 5820  
Db 5761 ACCTTCACTGCCCAACCCATATGATGCCCACTCGTATAGGAGTCTTTTAAAGGGAACCCC 5820  
Qy 5821 CATGCAATACTCATTTATGGACAGGAAATGATTAATCTAGTTGCTTGGAGACTT 5880  
Db 5821 CATGCAATACTCATTTATGGACAGGAAATGATTAATCTAGTTGCTTGGAGACTT 5880  
Qy 5881 GGAGTCACTGTCTGTGGACTTACTTCAACCAACTGTA TGTCTGATGGGGTGGAGTT 5940  
Db 5881 GGAGTCACTGTCTGTGGACTTACTTCAACCAACTGTA TGTCTGATGGGGTGGAGTT 5940  
Qy 5941 CAAGATCAGGCAAGAGAAAAA CATGTAAAAGAGTAATCTCCCAACTCAACCSGGGTACAT 6000  
Db 5941 CAAGATCAGGCAAGAGAAAAA CATGTAAAAGAGTAATCTCCCAACTCAACCSGGGTACAT 6000  
Qy 6001 GGCACTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCTCCGTACC 6060  
Db 6001 GGCACTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCTCCGTACC 6060  
Qy 6061 CATACTCGCTGGTAAGCTATTTAAATACCACTTCACTGGGCTCCATGAGGTCTCGGCC 6120  
Db 6061 CATACTCGCTGGTAAGCTATTTAAATACCACTTCACTGGGCTCCATGAGGTCTCGGCC 6120  
Qy 6121 CAAAACCTCACTAATCTGTGGATATGCTCCCTCTGAACTTCARGCCATATGTTTCAATC 6180  
Db 6121 CAAAACCTCACTAATCTGTGGATATGCTCCCTCTGAACTTCARGCCATATGTTTCAATC 6180  
Qy 6181 CTTGTACTGAAACATGGAACAACTTCAGCAAGAAATAAACACCACTTCGGTTTTAGTA 6240  
Db 6181 CTTGTACTGAAACATGGAACAACTTCAGCAAGAAATAAACACCACTTCGGTTTTAGTA 6240  
Qy 6241 GGACCTCTGTGTTTCCAAATSTGGAATAACCCATACCTCAAACTCACCTGTGTAAATTT 6300  
Db 6241 GGACCTCTGTGTTTCCAAATSTGGAATAACCCATACCTCAAACTCACCTGTGTAAATTT 6300  
Qy 6301 AGCAATACTATACAAACCAACTCCCAATGCGATCAGGTGGGTAACTCTCCCAACAAA 6360  
Db 6301 AGCAATACTATACAAACCAACTCCCAATGCGATCAGGTGGGTAACTCTCCCAACAAA 6360





PR 03-OCT-2000; 2000US-0237606P.  
 PR 03-OCT-2000; 2000US-0237608P.  
 PR 01-NOV-2000; 2000US-0244867P.  
 PR 01-NOV-2000; 2000US-0245084P.  
 XX (AVAL-) AVALON PHARM.  
 XX  
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
 PI Soppet DR, Weaver Z;  
 XX  
 XX WPI; 2002-188264/24.  
 XX  
 PT Screening for anti-neoplastic agent involves exposing cells to a chemical  
 PT agent to be tested for anti-neoplastic activity, and determining a change  
 PT in expression of a gene of a signature gene set.  
 XX  
 XX Claim 1; SEQ ID NO 81; 44pp; English.  
 PS  
 CC The present invention describes a method (M1) for screening for an anti-  
 CC neoplastic agent. The method involves exposing cells to a chemical agent  
 CC to be tested for anti-neoplastic activity, determining a change in  
 CC expression of at least one gene (I) of a signature gene set, where (I)  
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
 CC to ABL70110), or is at least 95% identical to (S), where a change in  
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
 CC activity and can be used in gene therapy. M1 can be used for screening an  
 CC anti-neoplastic agent, and can be used for producing a product which is  
 CC the data collected with respect to the anti-neoplastic agent as a result  
 CC of M1, and the data is sufficient to convey the chemical structure and/or  
 CC properties of the agent. M1 can be used in the treatment of cancer such  
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,  
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell  
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous  
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms  
 CC tumour  
 XX  
 SQ Sequence 56093 BP; 16164 A; 12346 C; 10702 G; 16881 T; 0 U; 0 Other;  
 Query Match 86.1%; Score 6448; DB 6; Length 56093;  
 Best Local Similarity 95.7%; Pred. No. 0;  
 Matches 6681; Conservative 195; Mismatches 50; Indels 53; Gaps 27;  
 QY 606 GATGGGAACGTTCCCGCAAGCAAAACGCGCTTAAGAGCTATTCTGGGAATTTGGGA 665  
 DB 30953 GATGGGAACGTTCCCGCAAGCAAAACGCGCTTAAGAGCTATTCTGGGAATTTGGGA 31012  
 QY 666 MCAATTTGACCTCAGACACTAAGAAAGAAACGACTTATTTCTGCGAGTGGCGCTG 725  
 DB 31013 CCAATTTGACCTCAGACACTAAGAAAGAAACGACTTATTTCTGCGAGTGGCGCTG 31072  
 QY 726 GCATCTCTGAGGAAGTATAAATTATAACACCATCTTAAGAGTACAGTCTTTTGTAGAA 785  
 DB 31073 GCATCTCTGAGGAAGTATAAATTATAACACCATCTTAAGAGTACAGTCTTTTGTAGAA 31132  
 QY 786 AAGGCAATGGAGTGAAGTCCATAGTCAAACTTTCTTTTCAATTAAGAGCACTCAC 845  
 DB 31133 AAGGCAATGGAGTGAAGTCCATAGTCAAACTTTCTTTTCAATTAAGAGCACTCAC 31192  
 QY 846 AATTATGTAAGAGTGTGATTATGCTTACAGGAGGCTTACAGAGTCTACCTCCCTATC 905  
 DB 31193 AATTATGTAAGAGTGTGATTATGCTTACAGGAGGCTTACAGAGTCTACCTCCCTATC 31252  
 QY 906 CCAGCATCCCGACTCTTCCCAATTAAGAGGACCCCTTCAACCCCAATGTTCCAA 965  
 DB 31253 CCAGCATCCCGACTCTTCCCAATTAAGAGGACCCCTTCAACCCCAATGTTCCAA 31312  
 QY 966 AAGGAGATAGCAAAAGGTAACAGTGAACCAAGAGTGCCTAATATCCCAATATGA 1025  
 DB 31313 AAGGAGATAGCAAAAGGTAACAGTGAACCAAGAGTGCCTAATATCCCAATATGA 31372  
 QY 1026 CCCCTCCCAAGCAGTGGGAGGAGAGAAATTCGGCCAGCCAGAGTGCATGTCYTTTTY 1085  
 DB 31373 CCCCT-CCAAGCAGTGGGAGGAGAGAAATTCGGCCAGCCAGAGTGCATGTCCTTTTC 31431

QY 1086 TCTCCAGACTTAAGCAAAATAAAAACAGACTTAGGTAAATTTCTCAGATAAATCCTGATGG 1145  
 DB 31432 TCTCCAGACTTAAGCAAAATAAAAACAGACTTAGGTAAATTTCTCAGATAAATCCTGATGG 31491  
 QY 1146 CTATATTGRTCTTTTACAAAGGTTAGGACAAATTTCTTGATCTGACATGGAGAGATATA 1205  
 DB 31492 CTATATTGRTCTTTTACAAAGGTTAGGACAAATTTCTTGATCTGACATGGAGAGATATA-A 31550  
 QY 1206 TGTCACTGCTAAATCAGACACTAAACCCCAATGAGAGAGTGCACCACTAACTGAGCCT 1265  
 DB 31551 TGTCACTGCTAAATCAGACACTAAACCCCAATGAGAGAGTGCACCACTAACTGAGCCT 31610  
 QY 1266 GAGGTTTGGCGATCTCTGATCTCAGTCAAGTCAATGGATANGATGACCAACAGAGG 1325  
 DB 31611 GAGGTTTGGCGATCTCTGATCTCAGTCAAGTCAATGGATANGATGACCAACAGAGG 31667  
 QY 1326 AAAGANAATGATTTCCCAAGCCAGCAGCAGTTCACAGTCTASACCTCATTTGGGAC 1385  
 DB 31668 AAAGANAATGATTTCCCAAGCCAGCAGCAGTTCACAGTCTAGACCTCATTTGGG--- 31724  
 QY 1386 ACAGAAATCAGTAACATGGGAGATTTGTCGAGACATTTGCTAACTTTGTCGTASAA 1445  
 DB 31725 ACAGAAATCAGTAACATTTGTCGAGACATTTGCTAACTTTGTCGTASAA 31783  
 QY 1446 GGACTAAGAAACTASGAGAAATCTAYGAATTAATCAATGATGTCACCAATTAACACA 1505  
 DB 31784 GGACTAAGAAACTTAGGAAG-AAAGTCTATGAATTAATCAATGATGTCACCAATTAACACA 31842  
 QY 1506 GGGGAAGGAAAGAAATCTCTACTGCTTTCTGGAGAGACTAAGGAGAGCATTTGAGGAAGC 1565  
 DB 31843 -GGGAGGAAAGAAATCTCTACTGCTTTCTGGAGAGACTAAGGAGAGCATTTGAGGAAGC 31901  
 QY 1566 GTGCTCTCTCTCAGTCACTGACTCTTCTGAAGGCCAACTAATCTTAAGCGTAAAGTTTATCA 1625  
 DB 31902 GTGCTCTCTCTCAGTCACTGACTCTTCTGAAGGCCAACTAATCTTAAGCGTAAAGTTTATCA 31961  
 QY 1626 CTGAGTCACTGTCAGACATTTAG-AAAGAACTTCAAGTCTGCGGTAGGCGCGAGCAAA 1684  
 DB 31962 CTGAGTCACTGTCAGACATTTAGAAAGAACTTCAAGTCTGCGGTAGGCGCGAGCAAA 32021  
 QY 1685 ACTTAGAAACCTATTGAACTTTGGCACTCGGTTTTTATAATAGAGATCAGGAGGAGC 1744  
 DB 32022 ACTTAGAAACCTATTGAACTTTGGCACTCGGTTTTTATAATAGAGATCAGGAGGAGC 32081  
 QY 1745 AGGCGGAACAGGACAAACCGGATTAAGGAGGCAACCGCTTTAGTCAATGACCTCAGG 1804  
 DB 32082 AGGCGGAACAGGACAAACCGGATTAAGGAGGCAACCGCTTTAGTCAATGACCTCAGG 32141  
 QY 1805 CAAGTGGAGCTTTGGAGGCTCTGGAAGGGAAGGAGTGGGCAATTTGAATGCTAATAGG 1864  
 DB 32142 CAAGTGGAGCTTTGGAGGCTCTGGAAGGGAAGGAGTGGGCAATTTGAATGCTAATAGG 32201  
 QY 1865 GCTTGTCTCCAGTGGGCTCTCAAGGACACTTTTAAAAAGATTTGTCAGATAGAAAGTAAAG 1924  
 DB 32202 GCTTGTCTCCAGTGGGCTCTCAAGGACACTTTTAAAAAGATTTGTCAGATAGAAAGTAAAG 32261  
 QY 1925 CGGCGCTTCTGTCATGCGCTTTATTTCAAGGGAATCACTGGAAGGCGCCACTGCCCCAGG 1984  
 DB 32262 CGGCGCTTCTGTCATGCGCTTTATTTCAAGGGAATCACTGGAAGGCGCCACTGCCCCAGG 32321  
 QY 1985 GGACAAAGGCTTTTGTAGTCAAGACCACTTAACAGATGATCCAGCAGCAGACTGAGGG 2044  
 DB 32322 GGACAAAGGCTTCTCTGAGTCAAGACCACTTAACAGATGATCCAGCAGCAGACTGAGGG 32381  
 QY 2045 TGCTTGGGCGAGGCGCATCCCATGCTCAGGAGCCCTGGGTATGCTTGACC 2104  
 DB 32382 TGCTTGGGCGAGGCGCATCCCATGCTCAGGAGCCCTGGGTATGCTTGACC 32441  
 QY 2105 ATTGAGGCCAGGAGGTTGTTCTCTGGAACACTGGTGGCGTCTTTAGTCTTACTCTTTC 2164  
 DB 32442 ATTGAGGCCAGG-AGGTTGTTCTCTGGAACACTGGTGGCGTCTTTAGTCTTACTCTTTC 32500

QY	2165	TGTC	CCGGACAAC	TGCTCCTCCAGATCTGTCACTATTCTCAGGGGGTCCNTAAGACGGCA	2224
Db	32501	TGTC	CCGGACAAC	CTGTCTCCAGATCTGTCACTA-TCTCAGGGGGTCC-TAAGACGGCA	32558
QY	2225	GTCA	TAGATAC	TTTATCCAGCCACTAAGTTATAGCACTGGGAGCTTTATCTTTTCAC	2284
Db	32559	GTCA	TAGATAC	-TTTCCCGAGCCACTAAGTTATG-AC TGGGGAGCTTTATCTTTTCAC	32616
QY	2285	ATGCT	TTTCTAA	TTATGCTTTGAAGCCCACTACCTTGTATGGGAGAGACATCTTAGCAA	2344
Db	32617	ATGCT	TTTCTAA	TTATGCTTTGAAGCCCACTACCTTGTATGGGAGAGACATCTTAGCAA	32676
QY	2345	AAGC	AGGGGCATATATAC	CTGTAACATATAGAGAGAGAACACCCGGTTTGTGTGTTNCCCTG	2404
Db	32677	AAGC	AGGGGCATATATAC	CTGTAACATATAGAGAGAGAACACCCGGTTTGTGTGTTNCCCTG	32735
QY	2405	CTTG	AGGAAGCAATTAATCT	GAAAGTCTGGCAACAGAGAGACAATATGAGAGCCAA	2464
Db	32736	CTTG	AGGAAGCAATTAATCT	GAAAGTCTGGCAACAGAGAGACAATATGAGAGCGAG-CAAA	32794
QY	2465	GAAT	GCCGCTCTGTTTCAAG	TTAAACTAAAGGATTCACCTTCTCTTCCCTACCAAAGCA	2524
Db	32795	GAAT	GCCGCTCTGTTTCAAG	TTAAACTAAAGGATTCACCTTCTCTTCCCTACCAAAGCA	32854
QY	2525	GTAC	CCCTCAGACCCAGGCCCA	ACAAGGATTCGAAAGATTTGAAGGATCTTAAAGC	2584
Db	32855	GTAC	CCCTCAGACCCAGGCCCA	ACAAGGATTCGAAAGATTTGAAGGATCTTAAAGC	32914
QY	2585	CCAAG	GCTTAGTAAACCATGAT	CACTCCCTGCAATTCGTTAGTGGATTCAGGAGG	2644
Db	32915	CCAAG	GCTTAGTAAACCATGAT	CACTCCCTGCAATTCGTTAGTGGATTCAGGAGG	32967
QY	2645	CACAGA	AACCCAGTGGACAGTGGAGGGT	TAGTGCAGATCTCAGGATTTCAATGGAGGC	2704
Db	32968	TACAGA	AACCCACACAGACAGTGA	-GGTTAGTGCAGATCTCAGGATTTCAAT-GAGGC	33025
QY	2705	CGTT	GCCTTTTATACCCAGCTGTAC	TAGCCCTTATCTGTGMYTTCGCAATACAGA	2764
Db	33026	TGTT	TGTCCTATATAGCAGCTGTAC	TAGCCCTTATCTGTGMYTTCGCAATACAGA	33085
QY	2765	GGAAG	CAGAGTGGTTTACASTCT	TGACCTTGAAGTGCCTTCTCTGATCCCTGTACA	2824
Db	33086	GGAAG	CAGAGTGGTTTACAGTCT	TGACCTTGAAGTGCCTTCTCTGATCCCTGTACA	33145
QY	2825	TCCT	GACTCTCAATTTCTGTTTG	CCCTTTGAAGATCTTCAAACCCARCATCTCAACTCAC	2884
Db	33146	TCCT	GACTCTCAATTTCTGTTTG	CCCTTTGAAGATCTTCAAACCCARCATCTCAACTCAC	33205
QY	2885	CTGG	ACTRTTTTACC	CAAGGTTCAAGGATAGYCCCCTATTTGTCGCAAGGATTTAGC	2944
Db	33206	CTGG	ACTRTTTTACC	CAAGGTTTCAAGGATAGYCCCCTATTTGTCGCAAGGATTTAGC	33265
QY	2945	CCAAG	ACTTGAGCATYMTCATAC	CTGAGACTCTTCTCTCTCTGATGAGTGAATTT	3004
Db	33266	CCAAG	ACTTGAGCCAACTCTCAT	CTGAGACA--CTTGTCTTCGGTAGTGGATGATTT	33323
QY	3005	ACTTT	TRGVGCYRFTTCAGAAA	CCCTTGCCATCAAGCCACCAAGGCTCTTTMAATTT	3064
Db	33324	ACTTT	TRGVGCYRFTTCAGAAA	CCCTTGCCATCAAGCCACCAAGGCTCTTTMAATTT	33383
QY	3065	CCTG	CGVAC	CTGTGGCTACAMGGTTTCAAACCAPARGCTCACTCTGCTCACAGCAGGT	3124
Db	33384	CCTG	CGCTACTCTGTGGCTAC	TGTTTCCAAACCAAGGCTCAACTCTGCTCACAGCAGGT	33443
QY	3125	TAAAT	ACTTTAGGCTTAABATTTAT	CCAAAGGCAACAGGCCCTCAGTAGGGAAYATCCA	3184
Db	33444	-----TACT	TAGGGCTATAAATTTAT	CCAAAGGCAACAGGCCCTCAGTAGGGAACACATCCA	33499
QY	3185	GCCT	TACTGGCTTATCC	TCATCYCAAAACCCCTAAAGCAACTAAGRBRRTTCCCTGGCCT	3244
Db	33500	GCCT	TACTGGCTTATCC	TCATCCCAAAACCCCTAAAGCAACTAAGGGGATTTCCCTTGGCT	33559
QY	3245	AAVAG	GYTTCTGCGCAAWATGGAT	TCCCGAGTTTGGCRAAATATAGCCAGGYCATATWATA	3304

Db	34623	AGGCAGGACCTCCATTAGAAATGCTTATAAAACAACCCCTAGTATATAGGTAATCCCCCTC	34682	Qy	5464	CARAGGACCTGAAAAATCCAGGAGACAACGCTAGTATTTCTGTGAACTCTAGAGGATT	5523
Qy	4385	CGGHAACCAAGCCCCAGTACTACAGCAGAGAAACAGAAATGGGAAACCTCAGAGG-CAG	4443	Db	35763	CAAGGAAACCTGAAAAATCCAGGAGACAACGCTAGTATTTCTGTGAACTCTAGAGGATT	35822
Db	34683	CGGHAACCAAGCCCCAGTACTACAGCAGAGAAACAGAAATGGGAAACCTCAGAGGACAG	34742	Qy	5524	TGCGCCTGCTCTTCAAAACAACAACAGGAGGAAAGTAACATAAAATCATAAATCCCCCATG	5583
Qy	4444	TTTTCTCCCTCGGACGGTTAGCCACATGAAGAAGGAAATACATTTTTCCTGCAACTAT	4503	Db	35823	TGCGCCTGCTCTTCAAAACAACAACAGGAGGAAAGTAACATAAAATCATAAAT-CCCCATG	35881
Db	34743	TTTTCTCCCTCGGACGGTTAGCCACATGAAGAAGGAAATACATTTTTCCTGCAACTAT	34802	Qy	5584	GSCTCCCTCTATCATATTTTCTCTKTASTGTTSTTTTACCCTSTTTTCACTCTCATCTGCA	5643
Qy	4504	CCAATGGAAATTAATAAAACCTTTCAATCAAAACCTTTCACTTAGGCATCGATAGCACCA	4563	Db	35882	GCCCTCCCTTATCATATTTTCTCTTACTGTTCTTTTACCCTCTTTTCACTCTCATCTGCA	35941
Db	34803	CCAATGGAAATTAATAAAACCTTTCAATCAAAACCTTTCACTTAGGCATCGATAGCACCA	34862	Qy	5644	CCCCCTCCATCGCCCTGTATGACCACTAGCTTCCCTTACACAGAGTCTTCTATGGGAATG	5703
Qy	4564	TCARATGGCAAAATCATATTATTAATGGAACGAGGCCCTTTTCAAAATCATAGCAATATKT	4623	Db	35942	CCCCCTCCATCGCCCTGTATGACCACTAGCTTCCCTTACCAAGAGTCTTCTATGGGAATG	36001
Db	34863	TCAGATGGCAAAATCATATTATTAATGGAACGAGGCCCTTTTCAAAATCATAGCAATAGT	34922	Qy	5704	CAGCGTCCCGGAAATATTTGATGCCCCATCGTATAGGAGTCTTTTAAAGGGAACCCCCACC	5763
Qy	4624	CAGGCGCTGTGAATGTGCCARAAATAATCCCTGCTCTTATGCTTATGCTTCCAGG	4683	Db	36002	CAGCGTCCCGGAAATATTTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACC	36061
Db	34923	CAGGCGCTGTGAATGTGCCAGAGAAATAATCCCTGCTCTTATGCTTATGCTTCCAGG	34982	Qy	5764	TTCACTGCCACACCCCATATGCCCCGCAACTGCTATCACTCTGCACTCTTTTGCATGAT	5823
Qy	4684	AAAAAARAAACAGGCCATTACCTGARAARACTGGCAACTGATTTTACCCACACAGCCC	4743	Db	36062	TTCACTGCCACACCCCATATGCCCCGCAACTGCTATCACTCTGCACTCTTTTGCATGAT	36121
Db	34983	AGAACAAAGAACAGGCCATTACCTGGAAGACTGGCAACTGATTTTACCCACAGCCC	35042	Qy	5824	GCAATACTCATTTATGGAAGGAAATGATTAATCTAGTTGCTCTGGAGGACTTGA	5883
Qy	4744	AAACCTCAGGGATTTTCACTACTAGTCTGGGTATATATTTTCAACGGGTTGGCABAG	4803	Db	36122	GCAATACTCATTTATGGAAGGAAATGATTAATCTAGTTGCTCTGGAGGACTTGA	36181
Db	35043	AAACCTCAGGGATTTTCACTACTAGTCTGGGTATATATTTTCAACGGGTTGGCABAG	35102	Qy	5884	GTCACTGCTCTGTGACTTACTTCAACCAAACTGGTATGTCTGATGGGGTGGAGTTCAA	5943
Qy	4804	GCTTTCCCTGTAGACAGAAAGGCCCAAGAGTAATAAAGGCACTAGTTTCATCAATA	4863	Db	36182	GTCACTGCTCTGTGACTTACTTCAACCAAACTGGTATGTCTGATGGGGTGGAGTTCAA	36241
Db	35103	GCTTTCCCTGTAGACAGAAAGGCCCAAGAGTAATAAAGGCACTAGTTTCATCAATA	35162	Qy	5944	GATCAGCAAGAGAAAAACATGTAAAAAGATTAATCTCCCAACTCACCSGGGTACATGCG	6003
Qy	4864	ATTCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCCAGGCC	4923	Db	36242	GATCAGCAAGAGAAAAACATGTAAAAAGATTAATCTCCCAACTCACCSGGGTACATGCG	36301
Db	35163	ATTCCAGATTCGGACTTCCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCCAGGCC	35222	Qy	6004	ACCTCTAGCCCTACAAAGGACTAGTCTCTCAAACTACATGAAACCTCCGTAACCCAT	6063
Qy	4924	ACAGTAACCCAGGGAGTATCCAGGGTTAGTATACGNATTCATTTACACTGCGCCTGA	4983	Db	36302	ACCTCTAGCCCTACAAAGGACTAGTCTCTCAAACTACATGAAACCTCCGTAACCCAT	36361
Db	35223	ACAGTAACCCAGGGAGTATCCAGGGCTTAGGTATACGNATTCATTTACACTGCGCCTGA	35282	Qy	6064	ACTGCGCTGTAAAGCTATTTAATAACCACTCCTACTGGGCTCCATGAGTCTCGGCCCAA	6123
Qy	4984	AGGCCACAGTCTCAGGAGAGTTCAGAGAAATGAATGAAYACTCAAGGACATCTAAA	5043	Db	36362	ACTGCGCTGTAAAGCTATTTAATAACCACTCCTACTGGGCTCCATGAGTCTCGGCCCAA	36421
Db	35283	AGGCCACAGTCTCAGGAGAGTTCAGAGAAATGAATGAAYACTCAAGGACATCTAAA	35342	Qy	6124	AACCTCACTAACTGTGGATATGCTCCCTGAACTTTCAAGGCCATATGTTTCAATCCCT	6183
Qy	5044	AAGCAAAACCCAGAAACCCACCTCACATGGCCTGTCTGTGCTTATAGCCCTTAAAAAGA	5103	Db	36422	AACCTCACTAACTGTGGATATGCTCCCTGAACTTTCAAGGCCATATGTTTCAATCCCT	36481
Db	35343	AAGCAAAACCCAGAAACCCACCTCACATGGCCTGTCTGTGCTTATAGCCCTTAAAAAGA	35402	Qy	6184	GTACTGTAACTGGAACAACCTTCAGCAGAGAAATAAACACCACTTCCGTTTACTAGGA	6243
Qy	5104	ATCTGCAACTTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCC	5163	Db	36482	GTACTGTAACTGGAACAACCTTCAGCAGAGAAATAAACACCACTTCCGTTTACTAGGA	36541
Db	35403	ATCTGCAACTTTTCCCAAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCC	35462	Qy	6244	CCTCTGTTTCCAAATSTGGAATTAACCCATACCTCAAACTCCTCCTGTGTAAAAATTAGC	6303
Qy	5164	TTCAATACCAATGACTTGTGCTTGAACCAAGCAGCCCACTTAGTTCGAGACATCACT	5223	Db	36542	CCTCTGTTTCCAAATSTGGAATTAACCCATACCTCAAACTCCTCCTGTGTAAAAATTAGC	36601
Db	35463	TTCAATACCAATGACTTGTGCTTGAACCAAGCAGCCCACTTAGTTCGAGACATCACT	35522	Qy	6304	AATACTACATACAAACCAACTCCCAATGCAATCAGGTGGGTAACTCTCCCTCCACACAAATA	6363
Qy	5224	CCTTAGCCAAAATATCAACAAGTTCTTAAAAATTAAGAAACCTTATCCCTCAGAGAGG	5283	Db	36602	AATACTACATACAAACCAACTCCCAATGCAATCAGGTGGGTAACTCTCCCTCCACACAAATA	36661
Db	35523	CCTTAGCCAAAATATCAACAAGTTCTTAAAAATTAAGAAACCTTATCCCTCAGAGAGG	35582	Qy	6364	GTCTGCCCTACCTCAGGAATATTTTGTCTGTGTGTAACCTCAGCCCTATCGTTTGTGAAT	6423
Qy	5284	GAAGAAGACTTATCCACCCWGTGACATGGTATAGTCAAGTCCCTTCCTTAATTC	5343	Db	36662	GTCTGCCCTACCTCAGGAATATTTTGTCTGTGTGTAACCTCAGCCCTATCGTTTGTGAAT	36721
Db	35583	GAAGAAGACTTATCCACCCCTTGTGACATGGTATAGTCAAGTCCCTTCCTTAATTC	35642	Qy	6424	GSCTCTTCAGAACTCATGTGCTTCTCTCATTCTTGTAGTGGCCCTCAATGRCATCTACACT	6483
Qy	5344	CATCCCTAGATACATCTGGAGAGACCTTACCCAGTCACTTTTATATACCCCACTGGG	5403	Db	36722	GSCTCTTCAGAACTCATGTGCTTCTCTCATTCTTGTAGTGGCCCTCAATGRCATCTACACT	36781
Db	35643	CATCCCTAGATACATCTGGAGAGACCTTACCCAGTCACTTTTATATACCCCACTGGG	35702	Qy	6484	GAACAAGATTTTATACAGTTATGTCTATCTAAGCCCGCCCAACAAAGAGTACCCATCTTT	6543
Qy	5404	TTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTTGAGTCAAAATCCTGGATCTGC	5463	Db	36782	GAACAAGATTTTATACAGTTATGTCTATCTAAGCCCGCCCAACAAAGAGTACCCATCTTT	36841
Db	35703	TTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTTGAGTCAAAATCCTGGATCTGC	35762				











Db 7523 CCTAGCCAAATATCAACAAGTTCTTAAAAANTTACAAGGAACCTATCCCTGAGAAGAGG 7582  
Qy 5284 GAAAGAAGACTATTCACCCWGTGACATGGTATTTAGTCAAGTCCCTTCYCTCTAATTCCTC 5343  
Db 7583 GAAAGAAGACTATTCACCCCTTGTGACATGGTATTTAGTCAAGTCCCTTCCTCTAATTCCTC 7642  
Qy 5344 CATCCCTAGATACATCTCTGGGAAGGACCTTACCAGTCAATTTATATACCCCAACTGCGG 5403  
Db 7643 CATCCCTAGATACATCTCTGGGAGAGCCCTTACCAGTCAATTTATATCTACCCCAACTGCGG 7702  
Qy 5404 TTAAGTGGCTGGAGTGGAGTCTTGGATACATCACATCTGAGTCAAAATCTCTGGATCTGC 5463  
Db 7703 TTAAGTGGCTGGAGTGGAGTCTTGGATACATCACATCTGAGTCAAAATCTCTGGATCTGC 7762  
Qy 5464 CAAAGGAACCTGAAATCTCAGAGACAAAGCTAGCTATTTCTGTGAAACCTCTAGAGGAT 5523  
Db 7763 CAAAGGAACCTGAAATCTCAGAGACAAAGCTAGCTATTTCTGTGTGAACCTCTAGAGGAT 7822  
Qy 5524 TGGCCTGCTCTTCAAAACAACACGAGGAGAAAGTAACTAAATCATAAATCCCCATG 5583  
Db 7823 TGGCCTGCTCTTCAAAACAACACGAGGAGAAAGTAACTAAATCATAAAT-CCCCATG 7881  
Qy 5584 GSCCTCCCTTATCAATTTTCTCTKTASTGTSTTTTACCCCTSTTTCACTCTCACTGCA 5643  
Db 7882 GSCCTCCCTTATCAATTTTCTCTTACTGTCTTTTACCCCTCTTCACTCTCACTGCA 7941  
Qy 5644 CCCCCTCCATGCGCTGTATGACACAGTAGCTCCCTCTYACWAGAGTTTCTATGGAGATG 5703  
Db 7942 CCCCCTCCATGCGCTGTATGACACAGTAGCTCCCTCTTACCAAGAGTTTCTATGGAGATG 8001  
Qy 5704 CAGCGTCCGGAAATATTGATGCCCATCGTATAGAGTCTTTSTAAGGGAACCCCCACC 5763  
Db 8002 CAGCGTCCGGAAATATTGATGCCCATCGTATAGAGTCTTTCTAAGGGAAACCCCCACC 8061  
Qy 5764 TTCACTGCCCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCAT 5823  
Db 8062 TTCACTGCCCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCATGCAT 8121  
Qy 5824 GCAATATCTCATTTATGGACAGGAAAAATGATTAATCTAGTTTCTGGAGGACTTGG 5883  
Db 8122 GCAATATCTCATTTATGGACAGGAAAAATGATTAATCTAGTTTCTGGAGGACTTGG 8181  
Qy 5884 GTCACTGTCTGTGGACTTACTTCAACCAACTGCTATGCTGTGAGGGGTGGAGTTCAA 5943  
Db 8182 GTCACTGTCTGTGGACTTACTTCAACCAACTGCTATGCTGTGAGGGGTGGAGTTCAA 8241  
Qy 5944 GATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCACTCAACCGGGTACATGCC 6003  
Db 8242 GATCAGGCAAGAGAAAAACATGTAAAAAGAGTAATCTCCCACTCAACCGGGTACATGCC 8301  
Qy 6004 ACCTTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCTCTCGTACCCAT 6063  
Db 8302 ACCTTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCTCTCGTACCCAT 8361  
Qy 6064 ACTCGCTGTGAAGCTATTTTAAATACCAACCTCACTGGGCTCCATGAGGTCTCGGCCAA 6123  
Db 8362 ACTCGCTGTGAAGCTATTTTAAATACCAACCTCACTGGGCTCCATGAGGTCTCGGCCAA 8421  
Qy 6124 AACCCTACTAACTGTGGATATGCTCTCCCTCGAACTTCARGCCATATGTTTCAATCCCT 6183  
Db 8422 AACCCTACTAACTGTGGATATGCTCTCCCTCGAACTTCARGCCATATGTTTCAATCCCT 8481  
Qy 6184 GTACTGAAACAATGGAACAATCTTACGACAGAGAAATAAACAACAATCTTCGCTTTTAGTGA 6243  
Db 8482 GTACTGAAACAATGGAACAATCTTACGACAGAGAAATAAACAACAATCTTCGCTTTTAGTGA 8541  
Qy 6244 CCTCTGTTTCCAAATSTGGAATACCCATACCTCAAACTCACCTGTGTAAATTTAGC 6303  
Db 8542 CCTCTGTTTCCAAATSTGGAATACCCATACCTCAAACTCACCTGTGTAAATTTAGC 8601  
Qy 6304 AATACTACATACAAACCAACTCCCAATGCACTCAGTGGGTAACTCTCCCAACACAAATA 6363  
Db 8602 AATACTACATACAAACCAACTCCCAATGCACTCAGTGGGTAACTCTCCCAACACAAATA 8661

Qy 6364 GTCTGCCCTACCTCAGGAATATTTTTTGTCTGTGTACCTCAGCCTATCGTTGTTTGAAT 6423  
Db 8662 GTCTGCCCTACCTCAGGAATATTTTTTGTCTGTGTACCTCAGCCTATCGTTGTTTGAAT 8721  
Qy 6424 GGCTCTTTCAGAACTATGTGTCTCTCTCATTTCTTAGTGCCCCYATGRCATCTACACT 6483  
Db 8722 GGCTCTTTCAGAACTATGTGTCTCTCTCATTTCTTAGTGCCCCCTATGACCATCTACACT 8781  
Qy 6484 GAAACAATTTATACAGTTATGTCAATCTTAAGCCCGCAACAAAAGAGTACCAATCTT 6543  
Db 8782 GAAACAATTTATACAGTTATGTCAATCTTAAGCCCGCAACAAAAGAGTACCAATCTT 8841  
Qy 6544 CCTTTTCTTAGAGCAGGAGTCTAGGTGCACTAGTACTGTCATTTGGCGGTATCAC 6603  
Db 8842 CCTTTTCTTAGAGCAGGAGTCTAGGTGCACTAGTACTGTCATTTGGCGGTATCAC 8901  
Qy 6604 ACCTCTACTCAGTTTCTACTACAACTATCTCAAGAACTTAAATGGGGACATCGAACGGGTC 6663  
Db 8902 ACCTCTACTCAGTTTCTACTACAACTATCTCAAGAACTTAAATGGGGACATCGAACGGGTC 8961  
Qy 6664 GCGACTCCCTGGTCACTTTGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTCA 6723  
Db 8962 GCGACTCCCTGGTCACTTTGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTCA 9021  
Qy 6724 AATCGAAGAGCTTTAGACTTGTAAACGCTGABAGGGGGAACCTGTTATTTTAGGG 6783  
Db 9022 AATCGAAGAGCTTTAGACTTGTAAACGCTGABAGGGGGAACCTGTTATTTTAGGG 9081  
Qy 6784 GAAGAATGCTGTTTATTTATGTTAACTCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 6843  
Db 9082 GAAGAATGCTGTTTATTTATGTTAACTCCGGAATCGTCACTGAGAAAGTTTAAAGAAAT 9141  
Qy 6844 CSAGATCGAATPACAACTGACAGAGAGCTTCGAAACACTGGAACCTCGGGGCTCTC 6903  
Db 9142 CGAGATCGAATPACAACTGACAGAGAGCTTCGAAACACTGGAACCTCGGGGCTCTC 9201  
Qy 6904 AGCCBATGATGCTCCCTGGATTCCTCCCTTCTTAGACCTCTAGCAGCTATTAATTTGTA 6963  
Db 9202 AGCCBATGATGCTCCCTGGATTCCTCCCTTCTTAGACCTCTAGCAGCTATTAATTTGTA 9261  
Qy 6964 CTCTCTTTGGACCTGTATCTTTTACCTCTGTTTAACTTTGTCTCTTCAGAAATCGAA 7023  
Db 9262 CTCTCTTTGGACCTGTATCTTTTAACTCTCTGTTTAACTTTGTCTCTTCAGAAATCGAA 9321  
Qy 7024 GCTGTAAAACTPACAAATGAGGCCCAAGATGACAGTCCCAAGACTAAGACTACCGCAGACC 7083  
Db 9322 GCTGTAAAACTPACAAATGAGGCCCAAGATGACAGTCCCAAGACTAAGACTACCGCAGACC 9381  
Qy 7084 CTGGACCGGCTGTGTAGCCACGATCTGATGTTAATGACATCAAGGGCACCCCTCCTGAG 7143  
Db 9382 CTGGACCGGCTGTGTAGCCACGATCTGATGTTAATGACATCAAGGGCACCCCTCCTGAG 9441  
Qy 7144 GAAATCTCAGCTGACAACTCTACTACGCCCAATTCAGCAGGAAGCAGTTAGAGCCGT 7203  
Db 9442 GAAATCTCAGCTGACAACTCTACTACGCCCAATTCAGCAGGAAGCAGTTAGAGCCGT 9501  
Qy 7204 SGTCCGCCAACTCCCAACAGCACTTAGGTTTTCTGTTGAGATGGGGGACTGAGAGAC 7263  
Db 9502 C-TCCGCCAACTCCCAACAGCACTTAGGTTTTCTGTTGAGATGGGGGACTGAGAGAC 9560  
Qy 7264 AGGACTAGCTGGATTTCTAGGCTGATTAAGAAATCCVTAAGCCTAGSTGGGAAGGTGACC 7323  
Db 9561 AGGACTAGCTGGATTTCTAGGCTGATTAAGAAATCCVTAAGCCTAGSTGGGAAGGTGACC 9620  
Qy 7324 ACATCCACTTTAAACACGGGCTTTGCAACTTAGVTCACTGACCTGACCAATCAGAGAGCTC 7383  
Db 9621 ACATCCACTTTAAACACGGGCTTTGCAACTTAGVTCACTGACCTGACCAATCAGAGAGCTC 9680  
Qy 7384 ACTAAAAATGCTTAATTTAGGCAAAAGACAGGAGTTAAAGAAATAGCCAAATCATTTATTGCMTG 7443  
Db 9681 ACTAAAAATGCTTAATTTAGGCAAAAGACAGGAGTTAAAGAAATAGCCAAATCATTTATTGCTGT 9740

QY 7444 AGACACAGCAGGAGGACATGATCGGATATAAACCCAGTTCGAGCGCGCAACGG 7503  
 Db |||||||  
 QY 9741 AGACACAGCAGGAGGACATGATCGGATATAAACCCAGTTCGAGCGCGCAACGG 9800  
 Db |||||||  
 QY 7504 CAACCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCACCTCT 7563  
 Db |||||||  
 QY 9801 CAACCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCACCTCT 9860  
 Db |||||||  
 QY 7564 ATTAATCTTGCARCTGCR 7582  
 Db |||||||  
 QY 9861 ATTAATCTTGCARCTGCA 9879  
 Db |||||||

RESULT 5  
 ID ADS30988 standard; DNA; 8523 BP.  
 XX  
 AC ADS30988;  
 XX  
 DT 18-NOV-2004 (first entry)  
 XX  
 DE Human genome high complexity repeat found in the HIRA gene #21.  
 XX  
 KW Human; ds;  
 KW histone cell cycle regulation defective, *S. cerevisiae* homologue A; HIRA;  
 KW high complexity repeat; in situ hybridisation; Southern blot;  
 KW chromosome breakpoint; inherited genetic disease; neoplastic disorder;  
 KW chromosome 22; DiGeorge syndrome; Velo-Cardio-facial syndrome.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003224356-A1.  
 XX  
 PD 04-DEC-2003.  
 XX  
 PF 14-MAY-2001; 2001US-00854867.  
 XX  
 PR 16-MAY-2000; 2000US-00573080.  
 XX  
 PA (KNOL/) KNOLL J H M.  
 PA (ROGA/) ROGAN P K.  
 PI Knoll JHM, Rogan PK;  
 XX  
 XX WPI; 2002-062378/08.  
 XX  
 PT Single copy genomic hybridization probes for detecting specific nucleic  
 PT acid sequences in sample by in situ hybridization useful for detection of  
 PT acquired or inherited genetic diseases.  
 XX  
 PS Example 1; SEQ ID NO 21; 30pp; English.  
 XX  
 CC The invention relates to a nucleic acid hybridisation probe comprising a  
 CC labelled, single copy nucleic acids of at least 50 nucleotides, which  
 CC will hybridise to a deduced single copy sequence interval in target  
 CC nucleic acid (TNA) of known sequence. The single copy sequence is deduced  
 CC by comparing the target nucleic acid (e.g. a disease causing gene) with a  
 CC collection of high and low complexity repeat sequences as found in the  
 CC genome of the organism from containing the target nucleic acid. The probe  
 CC is generated by PCR on the target sequence. The probe is essentially free  
 CC of blocking nucleic acid sequences which will hybridise to repeat  
 CC sequences within the genome of which the TNA is a part, and is labelled  
 CC with a label selected from fluorochrome-responsive labels, fluorochromes,  
 CC calorimetric chemical, conjugated proteins, antibodies, antigens and  
 CC their mixtures. The probe is useful in a hybridisation method, where the  
 CC hybridisation method is from in situ hybridisation, Southern blot, and  
 CC other methods in which nucleic acid is immobilised, where the method  
 CC further comprises selecting a single copy nucleic acid which will  
 CC hybridise to a duplicate or triplicon sequence domain. The probe is useful  
 CC for determining the existence of previously unknown repeat sequence  
 CC families in a genome. The method comprises reacting a labelled probe with  
 CC the genome, causing the probe to hybridise and ascertaining if the probe  
 CC hybridises to the genome at more than three preferably ten different

CC locations as a determination of new repeat sequence family, where the  
 CC determining step comprises selecting the single copy sequence from a  
 CC duplicon or triplicon sequence domain. The probe is useful for  
 CC determining a chromosome breakpoint and is useful in the fields for  
 CC cytogenetics and molecular genetics for determining the presence of  
 CC specific nucleic acid sequences in a sample of eukaryotic origin, e.g.  
 CC the probes may be used to analyse specific chromosomal locations by in  
 CC situ hybridisation as a detection of acquired or inherited genetic  
 CC diseases especially for detection of genetic or neoplastic disorders.  
 CC Unlike prior art techniques, the probe permits more precise chromosomal  
 CC breakpoint determinations by in situ hybridisation. The genomic sequence  
 CC comprising the human HIRA gene (histone cell cycle regulation defective,  
 CC *S. cerevisiae*, homologue A) was analysed for single copy sequence  
 CC intervals for use as probes of the invention. HIRA is located on  
 CC chromosome 22 as a duplicate, deletions of 1 copy lead to DiGeorge and  
 CC Velo-Cardio-facial syndromes. The present sequence is a high complexity  
 CC repeat found within the human genome used to analyse the HIRA gene for  
 CC repeat regions. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030224356.  
 XX  
 XX

QY 606 GATGGGAAAGCTTCCCGGCAAGACAAACCGCCCTTAAGACGTATTCTGGARAAATGGGA 665  
 Db |||||||  
 QY 1923 GATGGGAAAGCTTCCCGGCAAGACAAACCGCCCTTAAGACGTATTCTGGARAAATGGGA 1982  
 Db |||||||  
 QY 666 MCAATTTGACCTTCAGACACTAAGAAAGAAACGACTTATTCTTCGAGTCCGCGCTG 725  
 Db |||||||  
 QY 1983 CCAATTTGACCTTCAGACACTAAGAAAGAAACGACTTATTCTTCGAGTCCGCGCTG 2042  
 Db |||||||  
 QY 726 GCACCTCTGAGGAGTATAAATTATAACACCATCTTACAGCTAGACCTCTTTGTAGAA 785  
 Db |||||||  
 QY 2043 GCACCTCTGAGGAGTATAAATTATAACACCATCTTACAGCTAGACCTCTTTGTAGAA 2102  
 Db |||||||  
 QY 786 AAGGCAATGGAGTGAAGTGCCATAGTACAAACTTCTTTTCATTAAGAGACAACCTCAC 845  
 Db |||||||  
 QY 2103 AAGGCAATGGAGTGAAGTGCCATAGTACAAACTTCTTTTCATTAAGAGACAACCTCAC 2162  
 Db |||||||  
 QY 846 AATTATGTAAGTGTGATTTATGCCCTACAGGAAGCCCTTCAGAGTCTACCTCCCTATC 905  
 Db |||||||  
 QY 2163 AATTATGTAAGTGTGATTTATGCCCTACAGGAAGCCCTTCAGAGTCTACCTCCCTATC 2222  
 Db |||||||  
 QY 906 CCAGCATCCCGACTCTTCCCGCCTTAAGAGACCCCTTCAACCCAAATGGTCCAA 965  
 Db |||||||  
 QY 2223 CCAGCATCCCGACTCTTCCCGCCTTAAGAGACCCCTTCAACCCAAATGGTCCAA 2282  
 Db |||||||  
 QY 966 AAGGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTGCCAAATATTCGCCAAATATGA 1025  
 Db |||||||  
 QY 2283 AAGGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTGCCAAATATTCGCCAAATATGA 2342  
 Db |||||||  
 QY 1026 CCCCTCCCAAGCAGTGGGAGAGAGAAATTCGGCCCGCAGCCAGAGTGCATGTGCTTTT 1085  
 Db |||||||  
 QY 2343 CCCCT-CCAAGCAGTGGGAGAGAGAAATTCGGCCCGCAGCCAGAGTGCATGTGCTTTT 2401  
 Db |||||||  
 QY 1086 TCTCCAGACTTAAGCAAAATAAACAAGACTTAGGTAAATCTTCAGATAAATCTCATGG 1145  
 Db |||||||  
 QY 2402 TCTCCAGACTTAAGCAAAATAAACAAGACTTAGGTAAATCTTCAGATAAATCTCATGG 2461  
 Db |||||||  
 QY 1146 CTATATGTGTGTTTACAGGGTATAGGCAATCTTTGATCTTCAGATGAGAGATATATA 1205  
 Db |||||||  
 QY 2462 CTATATGTGTGTTTACAGGGTATAGGCAATCTTTGATCTTCAGATGAGAGATATATA-A 2520  
 Db |||||||  
 QY 1206 TGTCACTGCTAAATCAGACACTAAACCCCAAAATGAGAGAGTGCCACATAATCGAGCCT 1265  
 Db |||||||  
 QY 2521 TGTCACTGCTAAATCAGACACTAAACCCCAAAATGAGAGAGTGCCACATAATCGAGCCT 2580  
 Db |||||||  
 QY 1266 GAGRGTTGGCGATCTTGTTATCTCAGTCAGTCAATGGATANGSATGACCAAGG 1325  
 Db |||||||

Db 2581 GAGAGTTTTGGGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACAAACAG-AGG 2637  
Qy 1326 AAAGANAATGATTCCTCCACAGCCAGCAGCGAGTTCACAGTCTASACCTCATTTGGGGAC 1385  
Db 2638 AAAGAGAAATGATTCCTCCACAGCCAGCAGCGAGGTTCACAGTCTAGACCTCATTTGGG--- 2694  
Qy 1386 ACAGAAATCAGTAACATCGGAGATGGTGTCTGACAGACATTTGCTAACTTGTGTGTCTASAA 1445  
Db 2695 ACACAGANTCAGAACAT--GGAGATTTGGTGTCTGACAGACATTTGCTAACTTGTGTGTCTAGNA 2753  
Qy 1446 GGACTAAGGAAAACCTASGAAGAAATCTAYGAATTTACTCAATGATGTGCCACCATAAACACA 1505  
Db 2754 GGACTAAGGAAAACCTAGGAAG-AGTCTATGAATTTACTCAATGATGTCCACCATAAACACA 2812  
Qy 1506 GGGGAAGGAGAGAAAATCTTACTCTCTGAGAGAGACTAAGGAGGAGCATTTGAGGAAGC 1565  
Db 2813 -GGGAAGGAGAGAAAATCTTACTGCTCTCTGAGAGAGACTAAGGAGGAGCATTTGAGGAAGC 2871  
Qy 1566 GTGCCCTCTCTGTCACTGACTCTCTGTAAGGCCAACTAATCTTAAAGCGTAAAGTTTATCA 1625  
Db 2872 GTGCCCTCTCTGTCACTGACTCTCTGTAAGGCCAACTAATCTTAAAGCGTAAAGTTTATCA 2931  
Qy 1626 CTCACTCAGCTGACAGACATTAAG-AAAAAATTTCAAAAGTCTGCCGTAGGCCCGGAGCAAA 1684  
Db 2932 CTCACTCAGCTGACAGACATTAAGAAAAAATTTCAAAAGTCTGCCGTAGGCCCGGAGCAAA 2991  
Qy 1685 ACTTAGAAACCTTATTTGAACCTTGGCAACTTGGGTTTTTATTAATAGAGATCAGGAGGAGC 1744  
Db 2992 ACTTAGAAACCTTATTTGAACCTTGGCAACTTGGGTTTTTATTAATAGAGATCAGGAGGAGC 3051  
Qy 1745 AGGCGGAAACAGGACAAACGGGATTAAGAAAAAGGCCACCGCTTTAGTCAATGACCCCTCAGG 1804  
Db 3052 AGGCGGAAACAGGACAAACGGGATTAAGAAAAAGGCCACCGCTTTAGTCAATGACCCCTCAGG 3111  
Qy 1805 CAAAGTGGACTTTGGAGGCTCTGGAAAAAGGAAAAAGCTGGGCAAAATGAAATGCTTAATAGG 1864  
Db 3112 CAAAGTGGACTTTGGAGGCTCTGGAAAAAGGAAAAAGCTGGGCAAAATGAAATGCTTAATAGG 3171  
Qy 1865 GTTGTCTTCCAGTGGCGGTCTACAAGGACACTTTTAAAAAGATTTGCCAAGTAGAAGTAAAG 1924  
Db 3172 GTTGTCTTCCAGTGGCGGTCTACAAGGACACTTTTAAAAAGATTTGCCAAGTAGAAGTAAAG 3231  
Qy 1925 CCGCCCTCTGTCACATGCCCCCTTATTAAGGGATCACTGGAAGGCCCACTGCCCCAGG 1984  
Db 3232 CCGCCCTCTGTCACATGCCCCCTTATTAAGGGATCACTGGAAGGCCCACTGCCCCAGG 3291  
Qy 1985 GGACAAAGTCTTTTGAAGTCAGAGCCTAACAAGATGATCCAGCAGCAGGACTGAGGG 2044  
Db 3292 GGACAAAGTCTTCTGAGTCAGAGGCACTAACAAGATGATCCAGCAGCAGGACTGAGGG 3351  
Qy 2045 TGCTGGGGCAAGCGCCATGCCATGCCATCACCTCACAGAGCCCTGGGGATGCTTGACC 2104  
Db 3352 TGCTGGGGCAAGCGCCATGCCATGCCATCACCTCACAGAGCCCTGGGGATGCTTGACC 3411  
Qy 2105 ATTGAGGGCCAGAGAGTGTCTCTGAGCACTGGTGGGCTCTTCTAGTCTTACTCTTC 2164  
Db 3412 ATTGAGGGCCAGG-AGGTTGTCTCTGAGCACTGGTGGGCTCTTCTAGTCTTACTCTTC 3470  
Qy 2165 TGTCGGGCAACTGTCTCTCAGATCTGTCACATTTCTGAGGGGGTCCNTAAGACGGGCA 2224  
Db 3471 TGTCGGGCAACTGTCTCTCAGATCTGTCACATTTCTGAGGGGGTCC- TAAGACGGGCA 3528  
Qy 2225 GTCACTAGATCTTTTCCAGCCACTAAGATTATGAATGAACTGGGGAGCTTTATTTCTTTCAC 2284  
Db 3529 GTCACTAGATAC-TTCTCCAGCCACTAAGATTATG- ACTGGGGAGCTTTATTTCTTTCAC 3586  
Qy 2285 ATGCTTTCTAATATGCTTGAAGCCCACTACCTTGTAGGGAGAGACATTTCTAGGAA 2344  
Db 3587 ATGCTTTCTAATATGCTTGAAGCCCACTACCTTGTAGGGAGAGACATTTCTAGGAA 3646  
Qy 2345 AAGCAGGGGCCATTAATACACTGAAATAGGAGAGAAACACCCGCTTGTCTGTCNCCCTG 2404  
Db 3647 AAGCAGGGGCCATTAATACACTGAAATAGGAGAGAAACACCCGCTTGTCTGTCNCCCTG 3705

Qy 2405 CTTGAGGAGGAATTAATCTTGAAGTCTGGGCAACAGAGGACAAATATGGACGAGCCAAA 2464  
Db 3706 CTTGAGGAGGAATTAATCTTGAAGTCTGGGCAACAGAGGACAAATATGGACGAG-CAAA 3764  
Qy 2465 GAATGCCCGTCTCTTCAAGTTAAACTAAAGGATTTCCACTTTCTTCCCTACCAAAGGCA 2524  
Db 3765 GAATGCCCGTCTCTTCAAGTTAAACTAAAGGATTTCCAACCTCTTCTTCCCTACCAAAGGCA 3824  
Qy 2525 GTACCCCTCAGACCCAGGGCCAAACAGAGATTTCAAAAAGATTTGTTAAGGACTTAAAGC 2584  
Db 3825 GTA CCCCTCAGACCCAGGGCCAAACAGGACTCCAAAAGATTTGTTAAGGACCTTAAAGC 3884  
Qy 2585 CCAAGGCTTAGTAAACCATGCTAACTCCCTGAGATTAATTCCTGAGTGGATTTGAGGAGG 2644  
Db 3885 CCAAGGCTTAGTAAACCATGCTAACTCCCTGAGTAACTCCCTGAGTACTCC-----AATTTTAGAG 3937  
Qy 2645 CACAGAAACCCAGTGGGACAGTGGAGGGTTAGTCAAGATCTCAGGATTTATCAATGGAGGC 2704  
Db 3938 TACAGAAACCCAAACAGACAGTGGG-GGTTAGTCAAGATCTCAGGATTTATCAAT-GAGGC 3995  
Qy 2705 CGTTGTCTTTTATACCCAGCTGTACCTAGCCCTTATATCTGTGMYTTTCCCAAAATACAGA 2764  
Db 3996 TGTGTTCTCTATAGCCAGCTGTACCTAGCCCTTATATCTGTCTTCCCAAAATACAGA 4055  
Qy 2765 GGAAGCAGAGTGTTTACASTCCTGGACCTTMAAGTGCCTTCTCTGCATCCCTGTACA 2824  
Db 4056 GGAAGCAGAGTGTTTACAGTCCCTGGACCTTCAAGATGCTTCTGATCCCTGTACA 4115  
Qy 2825 TCCTGACTCTCAATCTTGTGCTTTGAAAGATCTTCAAAACCCARCATCTCAACTCAC 2884  
Db 4116 TCCTGACTCTCAATCTTGTGCTTTGAAAGATCTTCAAAACCCAAACATCTCAACTCAC 4175  
Qy 2885 CTGACCTRTTTTACCCCAAGGTTTCAAGGATGAGYCCCCCATCTA TTTTGGCCAGGCAATTAGC 2944  
Db 4176 CTGACTATTTTACCCCAAGGTTTCAAGGATGAGTCCCATCTA TTTTGGCCAGGCAATTAGC 4235  
Qy 2945 CCAAGACTTGAGYCARVTWTCATACCTGGACACTCTTGTCTCTGTAAGTGGATGATTT 3004  
Db 4236 CCAAGACTTGAGCCAACTCTCATCTCTGGA CA--CTTGTCTCTGTTAGGTTAGGATTT 4293  
Qy 3005 ACTTTTRCYGCCYRVTTCAGAAAACCTTGTGCATCAAGCCACCCAGGRCCTCTTMAATTT 3064  
Db 4294 ACTTTTGGCCGCCATTCAGAAAACCTTGTGCATCAAGCCACCCAGGCGCTTCAATTT 4353  
Qy 3065 CTTGCVACCTGTGGCTACAWGGTTTCAAAACSPARAGCTCARCTCTCTCACAGAGGT 3124  
Db 4354 CTTGCVACCTGTGGCTACATGTTTCAAAACCAAGGCTCAACTCTCTCACAGAGGT 4413  
Qy 3125 TAAATCTTAGGRCCTAATATTTATCCAAAGGACCCAGGCGCTCAGTGAAGGAYRATCCA 3184  
Db 4414 ----TACTTAGGGCTTAAATTTATCCAAAGGACCCAGGCGCTCAGTGAAGGAAACATCCA 4469  
Qy 3185 GCCTATACCTGGCTTATCTCTCATCYCAAAAACCTTAAAGCAACTAAGRGRRRTTCTTGGGRT 3244  
Db 4470 GCCTATACCTGGCTTATCTCTCATCCAAAACCTTAAAGCAACTAAGGGGATTTCTTGGCGT 4529  
Qy 3245 AAYAGGYTTTGGCCGAAWATGATTTCCCGAGTWTGGCAAATAGCCAGGYCATTAATA 3304  
Db 4530 AATAGGTTTCTGCCGAAAATGGAAT-CCAGGATAGGGGAAAATAGCCAGGTCATTAATA 4588  
Qy 3305 CASTAATTTAAGGAAACTCAGAAAGCCATACCCATTTARTAGATGAGYACTGAAGYMR 3364  
Db 4589 CACTAATTTAAGGAAACTCAGAAAGCCATACCCATTTTAGTAGATGGAACACTGAAGTAG 4648  
Qy 3365 AAGTGGCTTTCCAGGCCCTTAAAGAGGCTTTAAACCCCAAGYCCAGTGTTAAGYTTGCC 3424  
Db 4649 AAGTGGCTTTCCAGGCC-----TAACCCAGCCCAAGTGTTAAGYTTGCC 4694  
Qy 3425 AACRGGGCAAGACTTTTSTTATATYRTCAAGAAAAAACAAGAAAYAGCTCTRGGAGTCC 3484  
Db 4695 AACAGGGCAGACTTTTCTCATATGTACAG-AAAAAACAAGAAATAGCTCTTAGGAGTCC 4753

3485 TTACACAGTCCRAGGGAYGAGCTTGCAACCYRTGGCGYACCTGASTAAGCAAAATGATG 3544  
4754 TTACACAGATCCGAGGGATGAGCTTGCAACCTCTGGCATATCCCTGACTAAGGAAATGATG 4813  
3545 TAGTGGCAAAAGGGTGTGCTCATTTGTTTAYGGGTAGTGGTGGCAGTAGCAGTGYKTAGTAT 3604  
4814 TAGTGGCAAAAGGGTGTGACCTCATTTGTTTACGGGTAGTGGTGGCAGTAGCAGTCTTAGTAT 4873  
3605 CTGAAGCAGTTAAATAATACAGGAGATCTTACTGTGTGAGCATCTCATGAKGTGA 3664  
4874 CTGAAGCAGTTAAATAATACAGGAAAGAGATCTTACTGTGTGAGCATCTCATGATGTGA 4933  
3665 AYRGCTACTCACTGCTTAAAGGAGACTTGTGGCTGTGCAGCAACCYGTTTACTTAAATRTG 3724  
4934 ATGGCTACTCACTGCTTAAAGGAGACTTGTGGCTGTGCAGCAACCTGTTTACTTAAATGTC 4993  
3725 AGGCTCTATTATCTGAARGGCCAGTGTGCRACCTGTGCACTTGTGCAACTCTTAAACCCAG 3784  
4994 AGGCTCTATTATCTGAAGGGCCAGTGTGCGACTGTGCACTTGTGCAACTCTTAAACCCAG 5053  
3785 YCNCAATTTCTTCAGACAATGAAGAAAGATARAATAATTAATCTGTCAACAARTATTTCTC 3844  
5054 CCACATTTCTTCAGACAATGAAGAAAGATARAATAATTAATCTGTCAACAARTATTTCTC 5113  
3845 AAACCTATGCACTCTGAGGGGACCTTGTAGAGTTCCTTTGACTGTCCYTGACTCCYGGACCTTCAAC 3904  
5114 AAACCTATGCACTCTGAGGGGACCTTGTAGAGTTCCTTTGACTGTCCYTGACTCCYGGACCTTCAAC 5172  
3905 TTGTATCTGATGGAAGTTCCTTTGTAGAAAAGGACTTCGAAAAGYGGGGTATGCAGTG 3964  
5173 TTGTATCTGATGGAAGTTCCTTTGTAGAAAAGGACTTCGAAAAGYGGGGTATGCAGTG 5232  
3965 GTGAGTGATATGGAATATYTTGAAAGTAATCCCTCACTCCAGGAACCTAGTGTGYAGCTR 4024  
5233 GTGAGTGATATGGAATATCTTGAAGTAATCCCTCACTCCAGGAACCTAGTGTCCAGCTA 5292  
4025 GCAGAACTAATACCCCTCACTTGGGCACTAGAAATAGGAGAGAAAGGGYAAATATA 4084  
5293 GCAGAACTAATACCCCTCACTTGGGCACTAGAAATAGGAGAGAAAGGGYAAATATA 5352  
4085 TATACAGACTCTTARTATGCTYACCTAGTCTNTCCATGCCCATGMRCAATATGSRAGAA 4144  
5353 TATACAGACTCTTARTATGCTYACCTAGTCTNTCCATGCCCATGMRCAATATGSRAGAA 5412  
4145 AGGGAATTCCTAATCTCYGAGGAAACCTTATCAMACATCAGGAAGCCATAGGABATTA 4204  
5413 AGGGAATTCCTAATCTCYGAGGAAACCTTATCAMACATCAGGAAGCCATAGGABATTA 5472  
4205 TTATYTCGCHGTACAGAACTTACAGAGGTGGAGTCTTACACTGCTGGGTCTATCANAAA 4264  
5473 TTATYTCGCHGTACAGAACTTACAGAGGTGGAGTCTTACACTGCTGGGTCTATCANAAA 5532  
4265 GGAAAGAAAGGAAATASAGGAAATGCAAGCAKATATTGAAGCMAAAGAGCTGCA 4324  
5533 GGAAAGAAAGGAAATAGAGAAATGCAAGCAKATATTGAAGCMAAAGAGCTGCA 5592  
4325 AGGCAAGCCCTCCATTAGAAATGCTTATTAACCTTCCCTTAGTATAGGGTAATCCCTTC 4384  
5593 AGGCAAGCCCTCCATTAGAAATGCTTATTAACCTTCCCTTAGTATAGGGTAATCCCTTC 5652  
4385 CGGGAACCCAGCCAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCAGAGG-CAG 4443  
5653 CGGGAACCCAGCCAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCAGAGGACAG 5712  
4444 TTTTCTCCCTCGGGAGGTTAGCCACTGAGAGGGGAAATATCTTTTGCCTGCACTAT 4503  
5713 TTTTCTCCCTCGGGAGGTTAGCCACTGAGAGGGGAAATATCTTTTGCCTGCACTAT 5772  
4504 CCAATGGAATTAATCTTAAACCCCTTCACTCAAACTTTCTAGGATCGATAGCACCCA 4563  
5773 CCAATGGAATTAATCTTAAACCCCTTCACTCAAACTTTCTAGGATCGATAGCACCCA 5832  
4564 TCARATGGCCAAATCATTTATTTACTGGAACGAGCCCTTTTCAAACTATCAAGCARATAT 4623

5833 TCAGATGGCCAAATCATTTATCTGGACGAGCCCTTTTCAAACTATCAAGCAGATAGT 5892  
4624 CAGGGCTGTGAATCTGCGCARAAAATAATCCCTGCTTATCGCCCAAGCTCTCTCAGG 4683  
5893 CAGGGCTGTGAATCTGCGCARAAAATAATCCCTGCTTATCGCCCAAGCTCTCTCAGG 5952  
4684 ARBAACAAARAAACAGGCCATTTACCTGRRARAACTCGGCAACTGATTTTACCACAGGCC 4743  
5953 AGAACAAAGAACAGGCCATTTACCTGGAAGAGACTGGCAACTGATTTTACCACAGGCC 6012  
4744 AAACCTCAGGGATTTTCAGTATCTACTAGTCTGGGTARATATCTTTCAGGGTGGGCGARAG 4803  
6013 AAACCTCAGGGATTTTCAGTATCTACTAGTCTGGGTARATATCTTTCAGGGTGGGCGAG 6072  
4804 GCCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAAGGCACCTAGTTTCATGAATA 4863  
6073 GCCTTCCCTGTAGGACAGAAAAGGCCCAAGAGGTAAATAAAGGCACCTAGTTTCATGAATA 6132  
4864 ATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCC 4923  
6133 ATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCC 6192  
4924 ACAGTAACCCAGGAGTATCCAGGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCC 4983  
6193 ACAGTAACCCAGGAGTATCCAGGGCTTACAGAGTGACAATAGCCCTGCTTTCCAGGCC 6252  
4984 AGGCCACAGTCTCCTCAGGAAAGGTTCAGAAAATGAATGAAYACTCAAAGGCACATCTAAA 5043  
6253 AGGCCACAGTCTCCTCAGGAAAGGTTCAGAAAATGAATGAAYACTCAAAGGCACATCTAAA 6312  
5044 AAGCAAAACCCAGAAACCCCACTCATGCGCTGTCTGTGCTTATAGCCTTAAAGAA 5103  
6313 AAGCAAAACCCAGAAACCCCACTCATGCGCTGTCTGTGCTTATAGCCTTAAAGAA 6372  
5104 ATCTGCAACTTTTCCCAAAAGCAGACTTAGCCCATACGAAATGCTGTATGGAAGGCC 5163  
6373 ATCTGCAACTTTTCCCAAAAGCAGACTTAGCCCATACGAAATGCTGTATGGAAGGCC 6432  
5164 TTCTAATCAATCACTGTGCTGTGACCAAGACAGCAGCACTTGTGTGAGACATCACCT 5223  
6433 TTCTAATCAATCACTGTGCTGTGACCAAGACAGCAGCACTTGTGTGAGACATCACCT 6492  
5224 CTTAGCCAAATATCAACAAAGTTCTTAAACATTAACAGAACTTATCCCTGAGAAAGG 5283  
6493 CTTAGCCAAATATCAACAAAGTTCTTAAACATTAACAGAACTTATCCCTGAGAAAGG 6552  
5284 GAAAGAACTATTTCCACCWGTGATGTTAGTCAAGTCCCTTCTCTTAATCCC 5343  
6553 GAAAGAACTATTTCCACCWGTGATGTTAGTCAAGTCCCTTCTCTTAATCCC 6612  
5344 CATCCCTAGATACATCTGGGAAGGACCTTACCAGTCAATTTATTTATCCCAACTGCCG 5403  
6613 CATCCCTAGATACATCTGGGAAGGACCTTACCAGTCAATTTATTTATCCCAACTGCCG 6672  
5404 TTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACTTGGAGTCAAACTCTGATACTGC 5463  
6673 TTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACTTGGAGTCAAACTCTGATACTGC 6732  
5464 CAAAGAACTGTAATAATCAGGAGCAAGCTGATGTTTCTGTGAACCTCTTAGAGATT 5523  
6733 CAAAGAACTGTAATAATCAGGAGCAAGCTGATGTTTCTGTGAACCTCTTAGAGATT 6792  
5524 TCGGCTGCTCTTCAACAAACAGGAGGAGTAACTAAATCATAAATCCCCCATG 5583  
6793 TCGGCTGCTCTTCAACAAACAGGAGGAGTAACTAAATCATAAATCCCCCATG 6851  
5584 GSCCTCCCTTATCATATTTTCTKTASTGTTTCTTACCTCTTCTCTCTCTCTCTCTGCA 5643  
6852 GSCCTCCCTTATCATATTTTCTKTASTGTTTCTTACCTCTTCTCTCTCTCTCTCTGCA 6911  
5644 CCCCCTCCATGCGCTGTATGACAGTACCTCCCTTACCMAGAGTTCCTTATGGAGAAATG 5703



Db 6912 CCCCCCTCCATGCGCGTGTATGACCCAGTAGTCTCCCTTACCAGAGTTCCTTATGGAGATG 6971

Qy 5704 CAGCCTCCGGAAATATTGATGCCCATCGTATAGAGTCTTTTAAAGGAACCCCCACC 5763

Db 6972 CAGCGTCCGGAAATATTGATGCCCATCGTATAGAGTCTTTTAAAGGAACCCCCACC 7031

Qy 5764 TTCACGTGCCACACCACATATGCCCGCAACTGCTATCACTCTGCCACATCTTTTGCATGCAT 5823

Db 7032 TTCACGTGCCACACCACATATGCCCGCAACTGCTATCACTCTGCCACATCTTTTGCATGCAT 7091

Qy 5824 GCAAACTACTCAATATTGGACAGAGAAAAATGATTAATCTCTAGTTGTCTCTGGAGGACTTGGGA 5883

Db 7092 GCAAACTACTCAATATTGGACAGAGAAAAATGATTAATCTCTAGTTGTCTCTGGAGGACTTGGGA 7151

Qy 5884 GTCACTGTCTGTGGACTTACCTTCAACCAAACTGTGTATGTCTGATGGGGGTGGAGTTCAA 5943

Db 7152 GTCACTGTCTGTGGACTTACCTTCAACCAAACTGTGTATGTCTGATGGGGGTGGAGTTCAA 7211

Qy 5944 GATCAGGCAAGAGAAAAACATGTAAGAAGTAATCTCCCACTCACCGSGGTACATGCG 6003

Db 7212 GATCAGGCAAGAGAAAAACATGTAAGAAGTAATCTCCCACTCACCGSGGTACATGCG 7271

Qy 6004 ACCTCTAGCCCCACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCGTACCCAT 6063

Db 7272 ACCTCTAGCCCCACAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCGTACCCAT 7331

Qy 6064 ACTCGCTGGTAAGCCTATTAAATACACCCCTCACTGGGCTCCATGAGGTCTCGGCCCAA 6123

Db 7332 ACTCGCTGGTAAGCCTATTAAATACACCCCTCACTGGGCTCCATGAGGTCTCGGCCCAA 7391

Qy 6124 RACCCTACTACTGTGTGATATGCTCCCTGGAACCTTCARGCCATATGTTCAATCCCT 6183

Db 7392 RACCCTACTACTGTGTGATATGCTCCCTGGAACCTTCARGCCATATGTTCAATCCCT 7451

Qy 6184 GTACCTGAACAAATGGAACAACTTCAGCACAGAAATAAAACACACCTTCCTGTTTATAGGA 6243

Db 7452 GTACCTGAACAAATGGAACAACTTCAGCACAGAAATAAAACACACCTTCCTGTTTATAGGA 7511

Qy 6244 CCTCTGTTTCCAATSTGGAATAACCCATACCTCACTCAAACTCACTGTTGTAATAATTTAGC 6303

Db 7512 CCTCTGTTTCCAATSTGGAATAACCCATACCTCAAACTCACTGTTGTAATAATTTAGC 7571

Qy 6304 AATACTACATACACAAACCACTCCCAATGCATCAGGTGGTAACCTCCCCACACAAATA 6363

Db 7572 AATACTACATACACAAACCACTCCCAATGCATCAGGTGGTAACCTCCCCACACAAATA 7631

Qy 6364 GTCTGCCTACCCCTCAGGAATATTTTGTCTGTGTACTCTCAGCCCTATCGTTGTTGAAT 6423

Db 7632 GTCTGCCTACCCCTCAGGAATATTTTGTCTGTGTACTCTCAGCCCTATCGTTGTTGAAT 7691

Qy 6424 GGCTCTTCAGAAATCTATGTCTCTCTCATTTCTTAGTGCCCCCVATGCCATCTACACT 6483

Db 7692 GGCTCTTCAGAAATCTATGTCTCTCTCATTTCTTAGTGCCCCCVATGCCATCTACACT 7751

Qy 6484 GNAAGAATTTATACAGTTATGTCTATCTAAGCCCGCAACAAAGAGTACCCATTTCT 6543

Db 7752 GNAAGAATTTATACAGTTATGTCTATCTAAGCCCGCAACAAAGAGTACCCATTTCT 7811

Qy 6544 CTTTTTGTATAGGAGCAGGAGTGTAGTGTACATAGGTACTGTGCATTTGGCGGTATACA 6603

Db 7812 CTTTTTGTATAGGAGCAGGAGTGTAGTGTACATAGGTACTGTGCATTTGGCGGTATACA 7871

Qy 6604 ACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAAACGGGTC 6663

Db 7872 ACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAAACGGGTC 7931

Qy 6664 GCGGACTCCCTGGTCACTTGCAGATCACTTAACCTCCCTAGCAGCAGTAGTCTTCRA 6723

Db 7932 GCGGACTCCCTGGTCACTTGCAGATCACTTAACCTCCCTAGCAGCAGTAGTCTTCRA 7991

Qy 6724 AATCGAAGAGCTTTAGACTGTCTAAACCGCTGAGAGGGGGAACTCTGTTTATTTTAGGG 6783

Db 7992 AATCGAAGAGCTTTAGACTGTCTAAACCGCTGAGAGGGGGAACTCTGTTTATTTTAGGG 8051

Qy 6784 GAGAACTCTGTTATTTATGTTTAATCAATCCGAATCGTCACTGAGAAAGTTTAAAGAAATT 6843

Db 8052 GAGAAATGCTGTTTATTTATGTTTAATCAATCCGAATCGTCACTGAGAAAGTTTAAAGAAATT 8111

Qy 6844 CSAGATCGAAATACAAACGTAKAGCAGAGAGCTTCGAAACACTTGGACCTCGGGCCCTCCTC 6903

Db 8112 CGAGATCGAATACAAACGTAGAGCAGAGAGCTTCGAAACACTTGGACCTCGGGCCCTCCTC 8171

Qy 6904 AGCCRAATGAGTGCCTCGAATCTCCCTTCTTAGGACCTCTAGCAGCTTATAATATTGCTA 6963

Db 8172 AGCCAAATGAGTGCCTCGAATCTCCCTTCTTAGGACCTCTAGCAGCTTATAATATTGCTA 8231

Qy 6964 CTCCTCTTTGGACCTGTATCTTTTACCTCTCTTGTAACTTTTGTCTCTTCCAGAAATCGAA 7023

Db 8232 CTCCTCTTTGGACCTGTATCTTTTAACTCTCTTGTAACTTTTGTCTCTTCCAGAAATCGAA 8291

Qy 7024 GCTGTAAACTTACAAATGGAGCCCAAGATGCAGTCCAAAGACTTAAGATCTACCGCAGACCC 7083

Db 8292 GCTGTAAACTTACAAATGGAGCCCAAGATGCAGTCCAAAGACTTAAGATCTACCGCAGACCC 8351

Qy 7084 CTGGACCGGCTGYTAGCCCAAGATCTCTGATTTAATGACATCAAAGGCACCCCTCCTGAG 7143

Db 8352 CTGGACCGGCTGYTAGCCCAAGATCTCTGATTTAATGACATCAAAGGCACCCCTCCTGAG 8411

Qy 7144 GAAATCTCAGTGCACAACTCTACTACGCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 7203

Db 8412 GAAATCTCAGTGCACAACTCTACTACGCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 8471

Qy 7204 SGTGGCCCAACTCCCAACAGCAGCTTAGGTTTCTGTTGAGATGGGGAC 7255

Db 8472 GGTGGCCCAACTCCCAACAGCAGCTTAGGTTTCTGTTGAGATGGGGAC 8523

RESULT 6

ADY36376 ID ADY36376 standard; DNA; 8523 BP.

XX AC ADY36376;

XX AC 05-MAY-2005 (first entry)

XX DF HIRA genomic fragment SEQ ID NO 21.

XX DE hybridization; DNA detection; neoplasm; genetic disorder; cytogenetics;

XX KW HIRA; ds.

XX KW Homo sapiens.

XX OS WO200188089-A2.

XX PN 22-NOV-2001.

XX PD 15-MAY-2001; 2001WO-US015674.

XX PF 16-MAY-2000; 2000US-00573080.

XX PR 14-MAY-2001; 2001US-00854867.

XX PS (CHIL-) CHILDREN'S MERCY HOSPITAL.

XX PI Knoll JHM, Rogan PK, Cazarro PM;

XX DR WPI; 2002-062378/08.

XX PT Single copy genomic hybridization probes for detecting specific nucleic acid sequences in sample by in situ hybridization useful for detection of acquired or inherited genetic diseases.

XX PS Example 1; SEQ ID NO 21; 67pp; English.

XX CC The invention describes a nucleic acid hybridization probe (I) comprising a labeled, single copy nucleic acids of at least 50 nucleotides, which will hybridize to a deduced single copy sequence interval in target



CC nucleic acid (TNA) of known sequence. (I) is useful in a hybridization  
 CC method which comprises preparing a reaction mixture comprising RNA and  
 CC (I) which hybridizes to TNA, and causing (I) to hybridize to TNA, where  
 CC the hybridization method is from in situ hybridization, Southern blot,  
 CC and other methods in which nucleic acid is immobilized, where the method  
 CC further comprises selecting a single copy nucleic acid which will  
 CC hybridize to a duplcon or triplicon sequence domain. (I) is useful for:  
 CC determining the existence of previously unknown repeat sequence families  
 CC in a genome; determining a chromosome breakpoint and in the fields of  
 CC cytogenetics and molecular genetics for determining the presence of  
 CC specific nucleic acid sequences in a sample of eukaryotic origin, e.g.  
 CC the probes may be used to analyze specific chromosomal locations by in  
 CC situ hybridization as a detection of acquired or inherited genetic  
 CC diseases especially for detection of genetic or neoplastic disorders.  
 CC Unlike prior art techniques, (I) permits more precise chromosomal  
 CC breakpoint determinations by in situ hybridization. Hybridization  
 CC techniques utilizing (I), have made it possible to obtain reliable,  
 CC easily detectable signals with relatively small probes. A readily  
 CC detectable signal was obtained with a probe on the order of 2 kb in  
 CC length, using fluorescent in situ hybridization (FISH) technology. This  
 CC sensitivity of (I) is improved compared to the prior art, because the  
 CC probes of (I) are homogeneous single copy sequences. However, smaller  
 CC amplified segments, each comprising non-repetitive sequences, may also be  
 CC used in combination as probes to achieve adequate signals for in situ  
 CC hybridization. Complex single copy probes that hybridize to duplicated or  
 CC triplicated targets can also increase hybridization signals. This  
 CC sequence represents a human HIRA genomic sequence that shows homology to  
 CC a known high-complexity repeat sequence family of the human genome and is  
 CC used in the creation of an HIRA gene probe.

XX  
 SQ Sequence 8523 BP; 2560 A; 2118 C; 1814 G; 2031 T; 0 U; 0 Other;

Query Match 81.8%; Score 6124.6; DB 7; Length 8523;

Best Local Similarity 95.7%; Pred. No. 0;

Matches 6363; Conservative 186; Mismatches 50; Indels 53; Gaps 27;

QY 606 GATGGCAACGTTCCCGCAAGACAAACGCGCCCTAAGACGTAATCTGGAAATTTGGGA 665  
 DB 1923 GATGGCAACGTTCCCGCAAGACAAACGCGCCCTAAGACGTAATCTGGAAATTTGGGA 1982  
 QY 666 MCAATTTGACCCCTCAGACACTAAGAAAGAACGACTTATATCTTCGCGAGTCCGCGCTG 725  
 DB 1983 CCAATTTGACCCCTCAGACACTAAGAAAGAACGACTTATATCTTCGCGAGTCCGCGCTG 2042  
 QY 726 GCACCTCTGAGGAAGTATAAATATAACACCATCTTACAGCTAGACCTCTTTGTAGAA 785  
 DB 2043 GCACCTCTGAGGAAGTATAAATATAACACCATCTTACAGCTAGACCTCTTTGTAGAA 2102  
 QY 786 AAGGCAAAATGGAGTGAAGTGCCTAAGTACAAACTTTCTTTTCAATTAAGAGACAACCTAC 845  
 DB 2103 AAGGCAAAATGGAGTGAAGTGCCTAAGTACAAACTTTCTTTTCAATTAAGAGACAACCTAC 2162  
 QY 846 AATTATGTAAGTGTGATTTATGCGCTACAGGAGCCCTTCAGAGTCTACCTCCCTATC 905  
 DB 2163 AATTATGTAAGTGTGATTTATGCGCTACAGGAGCCCTTCAGAGTCTACCTCCCTATC 2222  
 QY 906 CCAGCATCCCGACTCTTCCCACTAATAAGGACCCCGCTTCAACCCCAATTTGGTCCAA 965  
 DB 2223 CCAGCATCCCGACTCTTCCCACTAATAAGGACCCCGCTTCAACCCCAATTTGGTCCAA 2282  
 QY 966 AAGGAGATAGCAAAAGGGTAAACAGTGAACCAAGAGTGCCTAATTTCCCAATTTATGA 1025  
 DB 2283 AAGGAGATAGCAAAAGGGTAAACAGTGAACCAAGAGTGCCTAATTTCCCAATTTATGA 2342  
 QY 1026 CCCCTCCAGCAGTGGGAGGAGAAATTCGCGCCAGCCAGAGTGCATGTGCTTTT 1085  
 DB 2343 CCCCT-CCAAGCAGTGGGAGGAGAAATTCGCGCCAGCCAGAGTGCATGTGCTTTTTC 2401  
 QY 1086 TCTCCAGACTTAAAGCAATAAAGACAGACTTAGGTAAATTTCTCAGATAAATTCCTGATGG 1145  
 DB 2402 TCTCCAGACTTAAAGCAATAAAGACAGACTTAGGTAAATTTCTCAGATAAATTCCTGATGG 2461  
 QY 1146 CTATATGTGTTTTACAAGGGTTAGGACAAATTTCTTTTGTATCTGACATGGAGATATATA 1205

DB 2462 CTATATGTGTTTTACAAGGGTTAGGACAAATTTCTTTGTATCTGACATGGAGATATA-A 2520  
 QY 1206 TGTCACTGCTAAATCAGACACTAAACCCAAATCAGAGAAAGTGCACCATAACTGACGCT 1265  
 DB 2521 TGTCACTGCTAAATCAGACACTAAACCCAAATCAGAGAAAGTGCACCATAACTGACGCT 2580  
 QY 1266 GAGGTTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATGGATANGATGACAAAGAGG 1325  
 DB 2581 GAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTCAATG--ATAGGATGACAAAG-AGG 2637  
 QY 1326 AAGANAAATGATTTCCCCACAGCCAGCAGCAGTCTCCAGTCTTASACCTCATTTGGGAC 1385  
 DB 2638 AAGAGAAATGATTTCCCCACAGCCAGCAGCAGTCTCCAGTCTTASACCTCATTTGGG-- 2694  
 QY 1386 ACAGAAATCAGTAACTAGGAGATTTGGTCTGCGACATTTTGTAACTTTGTGTCTASAA 1445  
 DB 2695 ACAGAAATCAGACAT-GGAGATTTGGTCTGCGACATTTTGTAACTTTGTGTCTAGAA 2753  
 QY 1446 GGACTAAGGAAATCTAGGAAATCTATGAAATTAATCAATGATGTCCACCAATACACA 1505  
 DB 2754 GGACTAAGGAAATCTAGGAAATCTATGAAATTAATCAATGATGTCCACCAATACACA 2812  
 QY 1506 GGGGAGGAGAAATCTTCTGCTTTCTGGAGACTTAAGGAGGCAATTCAGAGAGC 1565  
 DB 2813 -GGGAGGAGGAGAAATCTTCTGCTTTCTGGAGACTTAAGGAGGCAATTCAGAGAGC 2871  
 QY 1566 GTGCGCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAACTTTATCA 1625  
 DB 2872 GTGCGCTCTCTGTCACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGCGTAACTTTATCA 2931  
 QY 1626 CTCAGTCAGCTCAGACATTTAG-AAAAAATCTTCAAAAGTCTGCGTAGGCCCGGAGCAAA 1684  
 DB 2932 CTCAGTCAGCTCAGACATTTAGAAAAAATCTTCAAAAGTCTGCGTAGGCCCGGAGCAAA 2991  
 QY 1685 ACTTAGAAACCTTATTAATTTGGGCAACCTTCGGTTTTTTTATTAATAGAGATCAGAGAGC 1744  
 DB 2992 ACTTAGAAACCTTATTAATTTGGGCAACCTTCGGTTTTTTTATTAATAGAGATCAGAGAGC 3051  
 QY 1745 AGGCGGACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCCCTCAGG 1804  
 DB 3052 AGGCGGACAGGACAAACGGGATTTAAAAAAGGCCACCGCTTTAGTCATGACCCCTCAGG 3111  
 QY 1805 CAAGTGGACTTTGGAGGCTCTGGAAAAAGGAAAAAGCTGGGCAAAATTTGAATTCCTAATAGG 1864  
 DB 3112 CAAGTGGACTTTGGAGGCTCTGGAAAAAGGAAAAAGCTGGGCAAAATTTGAATTCCTAATAGG 3171  
 QY 1865 GCTTGTCTCAGTGGGCTCTACAAGGACACTTTTAAAAAAGATTTGTCAAAGTGAAGTAAG 1924  
 DB 3172 GCTTGTCTCAGTGGGCTCTACAAGGACACTTTTAAAAAAGATTTGTCAAAGTGAAGTAAG 3231  
 QY 1925 CCGCCCTTCTGTCATGCCCTTATTTCAAGGGAATCTCAAGGCGCCACTGCGCCCAAG 1984  
 DB 3232 CCGCCCTTCTGTCATGCCCTTATTTCAAGGGAATCTCAAGGCGCCACTGCGCCCAAG 3291  
 QY 1985 GGACAAAGTCTTTTGAAGTCAGAAAGCCTAATCAGAGATGATCCAGCAGCAGCTCAGG 2044  
 DB 3292 GGACAAAGTCTTTTGAAGTCAGAAAGCCTAATCAGAGATGATCCAGCAGCAGCTCAGG 3351  
 QY 2045 TGCCTGGGCAAGCGCCATGCCATCACCCTCACAAGAGCCCTGGGTATGCTTTGACC 2104  
 DB 3352 TGCCTGGGCAAGCGCCATGCCATCACCCTCACAAGAGCCCTGGGTATGCTTTGACC 3411  
 QY 2105 ATTGAGGCGCAGGAAGTGTCTCTGACACTGCTGGGTCTTCTTAGTCTTACTCTTC 2164  
 DB 3412 ATTGAGGCGCAGG-AGGTTGTCTCTGGACACTGCTGGGTCTTCTTAGTCTTACTCTTC 3470  
 QY 2165 TGTCCCGGACAACTGCTCTCAGACTGTCTACTTATTTCTGAGGGGTTCNTTAAGAGCGGCA 2224  
 DB 3471 TGTCCCGGACAACTGCTCTCAGACTGTCTACTTACTTACTTACTTACTTACTTACTT 3528  
 QY 2225 GTCACTAGATCTTTTCTCCAGGCACTAAGTTATGAAGTGGGAGCTTTATTTCTTTTCA 2284

Db 3529 GTCACTAGATAC-TTCTCCAGCCACTAAGTTATG-ACTGGGGAGCTTTTATTTCTTTTCAC 3586  
Qy 2285 ATGCTTTTCTAATTATGCTTGAAGCCCACTACCTTGTAGGGAGACATTTCTAGCAA 2344  
Db 3587 ATGCTTTTCTAATTATGCTTGAAGCCCACTACCTTGTAGGGAGACATTTCTAGCAA 3646  
Qy 2345 AAGCAGGGGCCAATTATACACCTTGAACATAGAGAGAGAACCCCGTTTGTGTNCCCTG 2404  
Db 3647 AAGCAGGGGCCAATTATACACCTTGAACATAGAGAGAGAACCCCGTTTGTGT-CCCTG 3705  
Qy 2405 CTTGAGGAAGGAATTAATCTCTGAAGTCTGGGCAACAGAGGACAATATGGACGACCAA 2464  
Db 3706 CTTGAGGAAGGAATTAATCTCTGAAGTCTGGGCAACAGAGGACAATATGGACGAG-CAAA 3764  
Qy 2465 GAATGCCCGTCTCTGTTCAAGTTAACTAAAGATTCACCTTCCCTTCCCTACCAAGGCA 2524  
Db 3765 GAATGCCCGTCTCTGTTCAAGTTAACTAAAGATTCACCTTCCCTTCCCTACCAAGGCA 3824  
Qy 2525 GTACCCCTCAGACCCCAAGGCCCAACAGGATTTCCAAAAGATTGTAAAGGACTTAAAGC 2584  
Db 3825 GTACCCCTCAGACCCCAAGGCCCAACAGGACTCCAAAAGATTGTAAAGGACTTAAAGC 3884  
Qy 2585 CCAAGGCTTAGTAAACCATGATTAATCTCCCTGAGTAAATCCGTTAGTGGATTGAGGAG 2644  
Db 3885 CCAAGGCTTAGTAAACCATGAGTAACCCCTGCACTCC-----AATTTTAGGAG 3937  
Qy 2645 CACAGAACCCAGTGGCAGTGGAGGGTTAGTGCAGATCTCAGGATTTCAATCGAGGC 2704  
Db 3938 TACAGAAACCCACAGACAGTGGG-GGTTAGTGCAGATCTCAGGATTTCAAT-GAGGC 3995  
Qy 2705 CGTGTCTTTTATACCAAGCTGTACTAGCCCTTATACGTGTMWTTCCCAATACCAGA 2764  
Db 3996 TGTGTCTCTATAGCAGCTGTACTAGCCCTTATACGTGTMWTTCCCAATACCAGA 4055  
Qy 2765 GGAAGCAGAGTGGTTTACASTCTCGACCTTMAGGATGCTTCTCTGTCATCCCTGTACA 2824  
Db 4056 GGAAGCAGAGTGGTTTACAGTCTCGACCTTCAGGATGCTTCTCTGTCATCCCTGTACA 4115  
Qy 2825 TCCTGACTCTCAATTTCTGTGTGGCTTTGAAGATCTTCAAAACCCAGACATCTCAACTCAC 2884  
Db 4116 TCCTGACTCTCAATTTCTGTGTGGCTTTGAAGATCTTCAAAACCCAGACATCTCAACTCAC 4175  
Qy 2885 CTGGAATTTTACCCAGGGTTCAGGGATAGYCCCATCTATTCTTGCCAGGCATTAGC 2944  
Db 4176 CTGGAATTTTACCCAGGGTTCAGGGATAGTCCCATCTATTCTTGCCAGGCATTAGC 4235  
Qy 2945 CCAAGACTTGAGCARTYMTACACTGGACACTCTTGTCCTTCRGTAGTGGATGATTT 3004  
Db 4236 CCAAGACTTGAGCARTYMTACACTGGACACTCTTGTCCTTCRGTAGTGGATGATTT 4293  
Qy 3005 ACTTTTGGYGCYRFTTCAGAAACCTTGTGCCATCAGCCACCCAGRCCTCTTAAATTT 3064  
Db 4294 ACTTTTGGYGCYRFTTCAGAAACCTTGTGCCATCAGCCACCCAGRCCTCTTAAATTT 4353  
Qy 3065 CCTCGYACTGTGGCTACAWGTTTCCAAAACCAAGCTCARTCTGCTCAGCAGCT 3124  
Db 4354 CCTCGYACTGTGGCTACAWGTTTCCAAAACCAAGCTCARTCTGCTCAGCAGCT 4413  
Qy 3125 TAAATCTTAGRCCTAARATTTACAAAGGCCACCGGCTCAGTGAGGAAYRYATCCA 3184  
Db 4414 ----TACTTAGGCCTAANATTTCCAAAGGCACCGGCTCAGTGAGGAACATCCA 4469  
Qy 3185 GCCTATCTAGGCTTATCTCATCYCAAAACCCCTAAAGCAACTAAGRRRTTCTTGGGRT 3244  
Db 4470 GCCTATCTAGGCTTATCTCATCYCAAAACCCCTAAAGCAACTAAGRRRTTCTTGGGRT 4529  
Qy 3245 AATAGGTTCTCCGGAATGATTTCCCGAGTWTGGCBAATAGCCAGGCTCATTAATA 3304  
Db 4530 AATAGGTTCTCCGGAATGATTTCCCGAGTWTGGCBAATAGCCAGGCTCATTAATA 4588  
Qy 3305 CASTAATTAAGGAACTCAGAAAGCCCAATACCCATTTARTAGATGGAYAMCTGAAGYMR 3364  
Db 4589 CACTAATTAAGGAACTCAGAAAGCCCAATACCCATTTAGTAAGATGGACAACTGAAGTAG 4648

Qy 3365 AAGTGGCTTTCCAGGCCCTTAAAGAGGCGCTTAAACCCCAAGYCCAGGTGTTAAGYTTGCC 3424  
Db 4649 AAGTGGCTTTCCAGGCCCTTAAAGAGGCGCTTAAACCCCAAGYCCAGGTGTTAAGYTTGCC 4694  
Qy 3425 AACRGGGCAAGACTTTTSTTATATYRTTACAGAAAAAACAGAAAYAGCTCTRGAGTCC 3484  
Db 4695 AACAGGGCAAGACTTTTCTTCATATGTACAG-AAAAAACAGGAATAGCTCTAGGAGTCC 4753  
Qy 3485 TTACACAGRTCCRAGGGAYAGCTTGCNAACCYRTGGCRYACTGTAATAGGAAATGTATG 3544  
Db 4754 TTACACAGRTCCRAGGGAYAGCTTGCNAACCYRTGGCRYACTGTAATAGGAAATGTATG 4813  
Qy 3545 TAGTGGCAAAAGGGTTGRCYTCATTTTAYGSGTAGTGTGCGTACAGTACAGTCTTAGTAT 3604  
Db 4814 TAGTGGCAAAAGGGTTGRCYTCATTTTAYGSGTAGTGTGCGTACAGTACAGTCTTAGTAT 4873  
Qy 3605 CTGAAGCAGTTTAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAGTGA 3664  
Db 4874 CTGAAGCAGTTTAAATAATACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAGTGA 4933  
Qy 3665 AYRGATATCTCACTGTCTAAAGGAGACTTGTGGCTGTGAGACAA CYGTTTACTTAAATRTC 3724  
Db 4934 ATGGCATACTCACTGTCTAAAGGAGACTTGTGGCTGTGAGACAACTGTTTACTTAAATGTC 4993  
Qy 3725 AGGCTCTATTAATCTTGAAGGCCAGTGTGCRACCTGTGCACTTGTGCAACTCTTAAACCCAG 3784  
Db 4994 AGGCTCTATTAATCTTGAAGGCCAGTGTGCRACCTGTGCACTTGTGCAACTCTTAAACCCAG 5053  
Qy 3785 YCNCAATTTCTTCCAGACAAATGAAGAAAGATARAAYATAA CTGTCAACAAATTAATTTCTC 3844  
Db 5054 CCAATTTCTTCCAGACAAATGAAGAAAGATARAAYATAA CTGTCAACAAATTAATTTCTC 5113  
Qy 3845 AAACTATGCTCGAGGGGACCTTCTGTAGAGTTTCCTGTGATCTGATCCYGAACCTTCAAC 3904  
Db 5114 AAACTATGCTCGAGGGGACCTTCTGTAGAGTTTCCTGTGATCTGATCCYGAACCTTCAAC 5172  
Qy 3905 TTGTATAGTGTGAGGATTTCTTGTAGAAAGAGACTTCGAAAGYGGGGTATCGATG 3964  
Db 5173 TTGTATAGTGTGAGGATTTCTTGTAGAAAGAGACTTCGAAAGYGGGGTATCGATG 5232  
Qy 3965 GTCACTGATATGAATAATTTGAAAGTAACTCCCTCACTCCAGGAACTAGTCTYAGCTR 4024  
Db 5233 GTCACTGATATGAGTAACTTTGAAAGTAACTCCCTCACTCCAGGAACTAGTCTYAGCTR 5292  
Qy 4025 GCAGAACTTAAATAGCCCTCACTTGGGCACTAGAAATTAGGAGAAAGGAAATATA 4084  
Db 5293 GCAGAACTTAAATAGCCCTCACTTGGGCACTAGAAATTAGGAGAAAGGAAATATA 5352  
Qy 4085 TATACAGACTCTTATATGCTYACTAGTCTTCCATGCCATGMBGCAATATGSRAGAA 4144  
Db 5353 TATACAGACTCTTATATGCTYACTAGTCTTCCATGCCATGMBGCAATATGSRAGAA 5412  
Qy 4145 AGGGAATTTCTTAACTTTCYAGGAGCACTTATCMAACATCAGGAAAGCCATTAGGATTA 4204  
Db 5413 AGGGAATTTCTTAACTTTCYAGGAGCACTTATCMAACATCAGGAAAGCCATTAGGATTA 5472  
Qy 4205 TTATYGGCWGTACAGAAACCTTARAGAGTGTGAGTCTTACACTGCTGCGGGTCACTNAAA 4264  
Db 5473 TTATYGGCWGTACAGAAACCTTARAGAGTGTGAGTCTTACACTGCTGCGGGTCACTNAAA 5532  
Qy 4265 GGAAAGAAAGGGGAATASAGGAAATGCCCAAGCAKATATTTGAAGCAAAAGAGCTGCA 4324  
Db 5533 GGAAAGAAAGGGGAATASAGGAAATGCCCAAGCAKATATTTGAAGCAAAAGAGCTGCA 5592  
Qy 4325 AGGAGGACCTTCCATTAAGAAATGCTTATTAACCTTCCCTTAGTATAGGGAATCCCTTC 4384  
Db 5593 AGGAGGACCTTCCATTAAGAAATGCTTATTAACCTTCCCTTAGTATAGGGAATCCCTTC 5652  
Qy 4385 CGGGAAACCAAGCCCTCAGTACTCAGCAGGAGAAACAGAAATGGGGAACTTCAAGG- CAG 4443  
Db 5653 CGGGAAACCAAGCCCTCAGTACTCAGCAGGAGAAACAGAAATGGGGAACTTCAAGGACAG 5712

QY 4444 TTTTCTCCCTCGGACGGTTAGCCACTGAAGAGGGAATAATCTTTTGCCTGCAACTAT 4503  
DB 5713 TTTTCTCCCTCGGACGGTAGCCACTGAAGAGGGAATAATCTTTTGCCTGCAACTAT 5772  
QY 4504 CAAATGGAATTAATCTTAAACCCCTTCATCAAAACCTTTCACTTAGGCATCGATAGCACCCA 4563  
DB 5773 CCAATGGAATTAATCTTAAACCCCTTCATCAAAACCTTTCACTTAGGCATCGATAGCACCCA 5832  
QY 4564 TCARATGGCCAAATCAATTAATCTTACTGGACAGGCGCTTTTCAAAACCTATCAGACATATX 4623  
DB 5833 TCAGATGGCCAAATCAATTAATCTTACTGGACAGGCGCTTTTCAAAACCTATCAGACATATX 5892  
QY 4624 CAGGCGCTGTGAATGTGCCRARAATAATCCCTGCCTTATCGCAAGCTCCTTCAGS 4683  
DB 5893 CAGGCGCTGTGAATGTGCCRAGAATAATCCCTGCCTTATCGCAAGCTCCTTCAGS 5952  
QY 4684 ARAACAAACAGGCGCTTACCTGRARAARACTGGCAACTGATTTTACCCACAGCC 4743  
DB 5953 AGAACAAACAGGCGCTTACCTGGAGAAGACTGGCAACTGATTTTACCCACAGCC 6012  
QY 4744 ARAACCTCAGGATTTTCACTATCTACTAGTCTGGGTABATATCTTCAAGGCTTGGGAG 4803  
DB 6013 ARAACCTCAGGATTTTCACTATCTACTAGTCTGGGTABATATCTTCAAGGCTTGGGAG 6072  
QY 4804 GCCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCAATGAATA 4863  
DB 6073 GCCTTCCCTGTAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACTAGTTTCAATGAATA 6132  
QY 4864 ATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCAGGCG 4923  
DB 6133 ATTTCCAGATTCGGACTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCAGGCG 6192  
QY 4924 ACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCTGA 4983  
DB 6193 ACAGTAAACCCAGGAGTATCCAGGCGTTAGGTATACGATATCACTTACACTGCGCTGA 6252  
QY 4984 AGGCCACAGTCTCAGGAGGCTCGAAGAAATGAATGAAYACTCAAGGACATCTAAA 5043  
DB 6253 AGGCCACAGTCTCAGGAGGCTCGAAGAAATGAATGAAYACTCAAGGACATCTAAA 6312  
QY 5044 AAGCAACCCAGGAAACCCACCTCACAATGGCTGTCTGTGCTATAGCCCTTAAAGAA 5103  
DB 6313 AAGCAACCCAGGAAACCCACCTCACAATGGCTGTCTGTGCTATAGCCCTTAAAGAA 6372  
QY 5104 ATCTGCAACTTTCCCAAAAGAGGACTTAGCCCATAGCAATGCTGTATGGAAGGCC 5163  
DB 6373 ATCTGCAACTTTCCCAAAAGAGGACTTAGCCCATAGCAATGCTGTATGGAAGGCC 6432  
QY 5164 TTTCATACCAATGACTTGTGCTTGACCCCAAGACGCAACTTAGTTGAGACATCACT 5223  
DB 6433 TTTCATACCAATGACTTGTGCTTGACCCCAAGACGCAACTTAGTTGAGACATCACT 6492  
QY 5224 CTTTAGCCAAATATCAACAGTTCTTAAACATTAAGGAACCTATCCCTGAGAGAGG 5283  
DB 6493 CTTTAGCCAAATATCAACAGTTCTTAAACATTAAGGAACCTATCCCTGAGAGAGG 6552  
QY 5284 GAAAGAACTATTCACCCMWGTGACATGTTAGTCAAGTCCCTTCCTCTAATTC 5343  
DB 6553 GAAAGAACTATTCACCCMWGTGACATGTTAGTCAAGTCCCTTCCTCTAATTC 6612  
QY 5344 CATCCTAGATACATCTCGGAGGAGCCCTA CCGAGTCAATTTATVTAACCCCACTGGG 5403  
DB 6613 CATCCTAGATACATCTCGGAGGAGCCCTA CCGAGTCAATTTATVTAACCCCACTGGG 6672  
QY 5404 TTAAGTGGCTGGAGTCTTGATACATCACTTGAATCAATCTGATCAATCTGATCTGC 5463  
DB 6673 TTAAGTGGCTGGAGTCTTGATACATCACTTGAATCAATCTGATCAATCTGATCTGC 6732  
QY 5464 CAAAGGAACTGAAATCCAGGAGCAACGCTAGCTATTTCTGTGAACCTCTAGAGGATT 5523  
DB 6733 CAAAGGAACTGAAATCCAGGAGCAACGCTAGCTATTTCTGTGAACCTCTAGAGGATT 6792  
QY 5524 TGCGCTGCTCTTCAAAACCAACAGGAGGAAGTAATCAATAATCATAAATCCCCCATG 5583

DB 6793 TGCGCTGCTCTTCAAAACCAACAGGAGGAAGTAATCAATAATCATAAAT-CCCATG 6851  
QY 5584 GSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTTACCCCTSTTTTCACTCTCACTGCA 5643  
DB 6852 GSCCTCCCTTATCATATTTTTCTCTKTASTGTTSTTTTACCCCTSTTTTCACTCTCACTGCA 6911  
QY 5644 CCGCTCCATGCGGCTGTATGACAGTAGTCTCCCTTACCMAGAGTTTCTATGGAGATG 5703  
DB 6912 CCGCTCCATGCGGCTGTATGACAGTAGTCTCCCTTACCMAGAGTTTCTATGGAGATG 6971  
QY 5704 CAGCGTCCCGGAAATATTTGATGCCCCATCGTATAGGAGTCTTTSTAAAGGAGCCCCACC 5763  
DB 6972 CAGCGTCCCGGAAATATTTGATGCCCCATCGTATAGGAGTCTTTCTAAGGAGAACCCACC 7031  
QY 5764 TTTACTGCCACACCCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTTGATGCA 5823  
DB 7032 TTTACTGCCACACCCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTTGATGCA 7091  
QY 5824 GCAATATCTCATATTTGGAAGGAAATGATTAATCTCTAGTTGCTCTCGAGGACTTGA 5883  
DB 7092 GCAATATCTCATATTTGGAAGGAAATGATTAATCTCTAGTTGCTCTCGAGGACTTGA 7151  
QY 5884 GTCACTGCTGTGACTTACTTCAACCCCACTGCTATGCTGATGGGGGTGGAGTTCAA 5943  
DB 7152 GTCACTGCTGTGACTTACTTCAACCCCACTGCTATGCTGATGGGGGTGGAGTTCAA 7211  
QY 5944 GATCAGGCAAGGAAACATGTAAGAGTAATCTCCCAACTCAACGCGGTACATGCG 6003  
DB 7212 GATCAGGCAAGGAAACATGTAAGAGTAATCTCCCAACTCAACGCGGTACATGCG 7271  
QY 6004 ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCGGTACCCAT 6063  
DB 7272 ACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCGGTACCCAT 7331  
QY 6064 ACTCGCTGTGAAGCTATTTAATACGACCCCTCACTGGGCTCCATGAGGTCTCGGCCAA 6123  
DB 7332 ACTCGCTGTGAAGCTATTTAATACGACCCCTCACTGGGCTCCATGAGGTCTCGGCCAA 7391  
QY 6124 AACCCCTACTAGTTTGGATAGCTCCCGCTGAATCTTCARGCCATATGTTTCAATCCCT 6183  
DB 7392 AACCCCTACTAGTTTGGATAGCTCCCGCTGAATCTTCARGCCATATGTTTCAATCCCT 7451  
QY 6184 GTACTGAACCAATGGAACAACTTACGACAGAAATAAACACCACTTCGCTTTAGTAGGA 6243  
DB 7452 GTACTGAACCAATGGAACAACTTACGACAGAAATAAACACCACTTCGCTTTAGTAGGA 7511  
QY 6244 CCTCTTGTTCGAATSTGGAATAACCCATACCTCAACCTCACTGCTGTGTAAATTTAGC 6303  
DB 7512 CCTCTTGTTCGAATSTGGAATAACCCATACCTCAACCTCACTGCTGTGTAAATTTAGC 7571  
QY 6304 AATACTACATACAAACCAACTCCCAATGCACTCAGGTGGTAATCTCCCAACCAATA 6363  
DB 7572 AATACTACATACAAACCAACTCCCAATGCACTCAGGTGGTAATCTCCCAACCAATA 7631  
QY 6364 GTCTGCCCTACCCCTCAGGAATAATTTTTCTGTGTGCTACTCAGCCATATCGTTTGTGAAT 6423  
DB 7632 GTCTGCCCTACCCCTCAGGAATAATTTTTCTGTGTGCTACTCAGCCATATCGTTTGTGAAT 7691  
QY 6424 GGTCTTTCAGAACTATGTGCTTCTCTCATTTTGTAGTGCCCTTATGRCATCTACACT 6483  
DB 7692 GGTCTTTCAGAACTATGTGCTTCTCTCATTTTGTAGTGCCCTTATGACCATCTACACT 7751  
QY 6484 GAAACGAATTTTATACAGTTATGTCTATCTAAGCCCGGCAACAAAGAGTACCCATTCTT 6543  
DB 7752 GAAACGAATTTTATACAGTTATGTCTATCTAAGCCCGGCAACAAAGAGTACCCATTCTT 7811  
QY 6544 CCTTTTGTATAGGAGCAGGAGTGTAGTGTCACTAGTGTGCGGATTTGGGGGTATCACA 6603  
DB 7812 CCTTTTGTATAGGAGCAGGAGTGTAGTGTCACTAGTGTGCGGATTTGGGGGTATCACA 7871  
QY 6604 ACCTCTACTAGTTTCTACTACAACTATCTCAAGAACTAAATGGGGACATGGAGCGGTC 6663



Qy	659	ATTGGGACCAATTTGACCTTCAGACACTAA-GAAAGAAA	CGACTTATATTTCTTCTGCAGT	717
Db	3976	ATTGGGACCAATTTGACCTTCAGACACTAAAGAAAGAAA	CGACTTATATTTCTTCTGCAGT	4035
Qy	718	GCCGCTGGCACTCCTCAGGGAAGTATAAAATATAAA	CCATCTTACAGCTAGACATCTTT	777
Db	4036	GCCGCTGGCACTCCTCAGGGAAGTATAAAATATAAA	CCATCTTACAGCTAGACATCTTT	4095
Qy	778	TTGTAGAAAAGGCAAAATGGAGTGCCTTAAGTACAA	ATCTTTCTTTTCAATTAAGAGA	837
Db	4096	TTGTAGAAAAGGCAAAATGGAGTGCCTTAAGTACAA	ATCTTTCTTTTCAATTAAGAGA	4155
Qy	838	CAACTCACAATTATGTATAAAAGTGTGATTTATGSCC	-TACAGGAGCCCTTCAGAGTCT	894
Db	4156	CAACTCACAATTATGTATAAAAGTGTGATTTATGSCC	-TACAGGAGCCCTTCAGAGTCT	4215
Qy	895	ACCTCCC-TATCCAGCATCCCCGACTCTCTCCCCAM	YTAATAAGAGACCCCCCTTCAAC	952
Db	4216	ACCTCCCCTATCCCGACATCCCCGACTCTCTCCCC	AACTAATAAGAGACCCCCCTTCAAC	4275
Qy	953	CCAAATGGTCCAAAGGAGATAGACAAAAGGTTAAAC	GAGTGAACCAAGAGTGCCTATAT	1012
Db	4276	CCAAATGGTCCAAAGGAGATAGACAAAAGGTTAAAC	GAGTGAACCAAGAGTGCCTATAT	4335
Qy	1013	TCCCAATATAGCCCTCCCAAGCAGTGGGAGAGAGAA	TTCCGCCACAGCAGAGTGC	1072
Db	4336	TCCCAATATAGCCCTCCCAAGCAGTGGGAGAGAGAA	TTCCGCCACAGCAGAGTGC	4395
Qy	1073	ATGTCVTTTTYFCTCCGACCTTAAAGCAATAAAAA	CAGACTTAGGTAAATTTCTCAG	1132
Db	4396	ATGTCCTTTTTTCTCTCCGACCTTAAAGCAATAA	AAAAACAGACTTAGGTAAATTTCTCAG	4455
Qy	1133	ATAAYCTGATGGCTATATTGRTGTTTTACAAGGTT	TAGACAAATCTTTTGATCTGACAT	1192
Db	4456	ATAACCTGATGGCTATATTGRTGTTTTACAAGGTT	TAGACAAATCTTTTGATCTGACAT	4515
Qy	1193	GGAGATATATATGTCACCTCTAAATCAGACACTAA	CCCCCAATGAGAGAGTGCACC	1252
Db	4516	GGAGATATATATGTCACCTCTAAATCAGACACTAA	CCCCCAATGAGAGAGTGCACC	4575
Qy	1253	ATAACTGCAGCTGAGRTTTGGCGATCTCTGGTAT	CTCAGTCAGGTCAATGGATANGGA	1312
Db	4576	ATAACTGCAGCTGAGRTTTGGCGATCTCTGGTAT	CTCAGTCAGGTCAATGGATANGGA	4634
Qy	1313	TGACAAACAGAGGAAAGANATGATTTCCCAACAG	GCACGAGTTCCTCAGTCTASAC	1372
Db	4635	TGACAAACAGAGGAAAGAG-ATGATTTCCCAACAG	GCACGAGTTCCTCAGTCTAGAC	4693
Qy	1373	CCTCATTTGGGGACACAGAAATCAGTACATGGGAG	ATTGGTGTGCGACATTTGCTAAC	1432
Db	4694	CCTCATTTGGGGACACAGAAATCAGTACATGGGAG	ATTGGTGTGCGACATTTGCTAAC	4753
Qy	1433	TTGTGTCTASAGGACTAAGGAAAACTASGAAGAA	RTCTAYCAATTACTCAATGATGT	1492
Db	4754	TTGTGTCTCTAAGGACTAAGGAAAACTAAGGAAAA	CTACGAAATTACTCAATGATGT	4813
Qy	1493	GCACCTAAACACAGGGGAAGGAAAGAAATCTCTA	TGCGCTTTCTGGAGAGACTAAGGAG	1552
Db	4814	GCACCTAAACACAGGGGAAGGAAAGAAATCTCTA	TGCGCTTTCTGGAGAGACTAAGGAG	4873
Qy	1553	GCATTTAGGGAAGCGTGCTCTGTGTACCTGAC	TCTTCTGAAGGCCAACTTAATCTTAAG	1612
Db	4874	GCATTTAGGGAAGCGTGCTCTGTGTACCTGAC	TCTTCTGAAGGCCAACTTAATCTTAAG	4933
Qy	1613	CGTAAGTTTATCACTCAGTCAGCTGCAGACATTA	GAAAAAACTTCAAAAGTCTGCCGTAG	1672
Db	4934	CGTAAGTTTATCACTCAGTCAGCTGCAGACATTA	GAAAAAACTTCAAAAGTCTGCCGTAG	4993
Qy	1673	GCCCGGAGCAAACTTAGAAAAACCTTATGAAC	CTTGGCAACVTCGGTTTTTATAATAGAG	1732
Db	4994	GCCCGGAGCAAACTTAGAAAAACCTTATGAAC	CTTGGCAACVTCGGTTTTTATAATAGAG	5053

Qy	1733	ATCAGGAGGACGAGCGGAAACAGGACAAACGGGATTA	17992	ATCAGGAGGACGAGCGGAAACAGGACAAACGGGATTA
Db	5054	ATCAGGAGGACGAGCGGAAACAGGACAAACGGGATTA	5113	ATCAGGAGGACGAGCGGAAACAGGACAAACGGGATTA
Qy	1793	ATGACCTCTCAGGCAAGTGAGCTTTGGAGGCTCTG	1852	ATGACCTCTCAGGCAAGTGAGCTTTGGAGGCTCTG
Db	5114	ATGACCTCTCAGGCAAGTGAGCTTTGGAGGCTCTG	5173	ATGACCTCTCAGGCAAGTGAGCTTTGGAGGCTCTG
Qy	1853	ATGCTTAATAGGGTGCTTTCAGATGCGGTCTACAA	1912	ATGCTTAATAGGGTGCTTTCAGATGCGGTCTACAA
Db	5174	ATGCTTAATAGGGTGCTTTCAGATGCGGTCTACAA	5233	ATGCTTAATAGGGTGCTTTCAGATGCGGTCTACAA
Qy	1913	AGTAGAGTAAGCGGCCCTTCGTCATGCCCTTATTT	1972	AGTAGAGTAAGCGGCCCTTCGTCATGCCCTTATTT
Db	5234	AGTAGAGTAAGCGGCCCTTCGTCATGCCCTTATTT	5293	AGTAGAGTAAGCGGCCCTTCGTCATGCCCTTATTT
Qy	1973	CACATGCCCCAGGGGACAAAGGCTCTTTTGAGTC	2032	CACATGCCCCAGGGGACAAAGGCTCTTTTGAGTC
Db	5294	CACATGCCCCAGGGGACAAAGGCTCTTTTGAGTC	5353	CACATGCCCCAGGGGACAAAGGCTCTTTTGAGTC
Qy	2033	CAGGACTGAGGGTGCTTGGGGCAAGCGCCATGCC	2092	CAGGACTGAGGGTGCTTGGGGCAAGCGCCATGCC
Db	5354	CAGGACTGAGGGTGCTTGGGGCAAGCGCCATGCC	5413	CAGGACTGAGGGTGCTTGGGGCAAGCGCCATGCC
Qy	2093	GTATGCTTGACCAATGAGGGCCAGGAAGTTGTCT	2152	GTATGCTTGACCAATGAGGGCCAGGAAGTTGTCT
Db	5414	GTATGCTTGACCAATGAGGGCCAGGAAGTTGTCT	5473	GTATGCTTGACCAATGAGGGCCAGGAAGTTGTCT
Qy	2153	GTCTTACTCTTCGTCCCGGACAACTGTCTCCAG	2212	GTCTTACTCTTCGTCCCGGACAACTGTCTCCAG
Db	5474	GTCTTACTCTTCGTCCCGGACAACTGTCTCCAG	5533	GTCTTACTCTTCGTCCCGGACAACTGTCTCCAG
Qy	2213	NTAAGCGGSCAGTCACTAGATACTTTTCTCCAG	2272	NTAAGCGGSCAGTCACTAGATACTTTTCTCCAG
Db	5534	-TAAGCGGSCAGTCACTAGATACTTTTCTCCAG	5592	-TAAGCGGSCAGTCACTAGATACTTTTCTCCAG
Qy	2273	TATTCCTTTTCAATGCTTTTCTTAATTAATGCT	2332	TATTCCTTTTCAATGCTTTTCTTAATTAATGCT
Db	5593	TATTCCTTTTCAATGCTTTTCTTAATTAATGCT	5652	TATTCCTTTTCAATGCTTTTCTTAATTAATGCT
Qy	2333	ACATTCCTAGCAAAAGCAGGGGCCATTAACACT	2392	ACATTCCTAGCAAAAGCAGGGGCCATTAACACT
Db	5653	ACATTCCTAGCAAAAGCAGGGGCCATTAACACT	5712	ACATTCCTAGCAAAAGCAGGGGCCATTAACACT
Qy	2393	GTGTGTCCTCTGCTGAGGAAGGAATTAATCTCT	2452	GTGTGTCCTCTGCTGAGGAAGGAATTAATCTCT
Db	5713	GTGTGT-CCCTCTGCTGAGGAGGAATTAATCT	5771	GTGTGT-CCCTCTGCTGAGGAGGAATTAATCT
Qy	2453	GGACGAGCCAAAGAAATGCCCCGCTCTGTTCAA	2512	GGACGAGCCAAAGAAATGCCCCGCTCTGTTCAA
Db	5772	GGACGAGCCAAAGAAATGCCCCGCTCTGTTCAA	5831	GGACGAGCCAAAGAAATGCCCCGCTCTGTTCAA
Qy	2513	CTACCAAAAGCAGTACCCCTCAGACCCCAAGGC	2570	CTACCAAAAGCAGTACCCCTCAGACCCCAAGGC
Db	5832	CTACCAAAAGCAGTACCCCTCAGACCCCAAGGC	5891	CTACCAAAAGCAGTACCCCTCAGACCCCAAGGC
Qy	2571	AAGGACTTAAAGCCCAAGGCTTAGTAAAAACAT	2630	AAGGACTTAAAGCCCAAGGCTTAGTAAAAACAT
Db	5892	AAGGACTTAAAGCCCAAGGCTTAGTAAAAACAT	5951	AAGGACTTAAAGCCCAAGGCTTAGTAAAAACAT
Qy	2631	GTGGATTGAGGAGGACAGAAACCCAGTGGAGGT	2690	GTGGATTGAGGAGGACAGAAACCCAGTGGAGGT
Db	5952	GTGGATTGAGGAGGACAGAAACCCAGTGGAGGT	6011	GTGGATTGAGGAGGACAGAAACCCAGTGGAGGT
Qy	2691	TTATCAATGAGGCCGTTGTCCTTTTATACCAGCT	2750	TTATCAATGAGGCCGTTGTCCTTTTATACCAGCT
Db	6012	TTATCAATGAGGCCGTTGTCCTTTTATACCAGCT	6071	TTATCAATGAGGCCGTTGTCCTTTTATACCAGCT
Qy	2751	TCCCAAAATACGAGGAAGCAGAGTGGTTTACAS	2810	TCCCAAAATACGAGGAAGCAGAGTGGTTTACAS
Db	6072	TCCCAAAATACGAGGAAGCAGAGTGGTTTACAS	6131	TCCCAAAATACGAGGAAGCAGAGTGGTTTACAS
Qy	2811	TGCAATCCCTGTATCATCTGACTCTCAATTTCT	2870	TGCAATCCCTGTATCATCTGACTCTCAATTTCT



Db 6132 TGCATCCCTGACATCCTGACTCTCAATCTTGTTGGCTTTGAAGATACTTCAAAACCA 6191  
Qy 2871 ROATCTCAACTCACTGAGCTRTTTTATCCCAAGGGTTCAAGGATAGYCCCAATCTATTT 2930  
Db 6192 ACATCTCAACTCACCTGGACTGTTTATCCCAAGGGTTCAAGGATAGYCCCAATCTATTT 6251  
Qy 2931 GCGCAGGATAGCCCAAGACTGAGYCARVTMTCTACCTTGGACACTCTTGTCTTCRG 2990  
Db 6252 GCGCAGGATAGCCCAAGACTGAGCAATCCTCATACCTGGACA--CTTGTCTTCGG 6309  
Qy 2991 TAKGTGGATGATTTACTTTTTRGCGCCYRTTCAGAAAACCTTGTCATCAAGCCACCA 3050  
Db 6310 TAGTGGATGATTTACTTTTGGCCCACTTCAAGAAAACCTTGTCATCAAGCCACCA 6369  
Qy 3051 GCRCTCTMAATTTCTCCYACCTGTGGCTACWGGTTTCCAAACSAARAGCTCARCTC 3110  
Db 6370 GCGCTCTTCAATTTCTCGTACCTGTGGCTACATGGTTTCCAAACCAAGGCTCAACTC 6429  
Qy 3111 TGCTCAGACAGGTTAAATACCTTAGGRTAARATTTATCAAGGCAACAGGCCCTCAGT 3170  
Db 6430 TGCTCAGACAGGTT---TACTTAGGCTTAAATTTATCCAAAGGCAACAGGCCCTCAGT 6485  
Qy 3171 GAGGAAYRATCCAGCTATACCTGCTTATCTCATCYCAAAACCTTAAAGCAACTAAGR 3230  
Db 6486 GAGGAACCAATCCAGCTTATACCTGCTTATCTCATCCCAAAACCTTAAAGCAACTAAG 6545  
Qy 3231 GRRTTCTTGGRTAAYAGGYTTCTGCCGAATWGGATTTCCCAAGTTWGGCAATAGC 3290  
Db 6546 GGATTTCTTGGGTAATAGGTTTCTGCCGAATWGGATTTCCCAAGTTWGGCAATAGC 6604  
Qy 3291 CAGGYCATTAATACASTAATTAAGGAACTCAGAAAGCAATACCCATTTARTTAAGATG 3350  
Db 6605 CAGGTCATTAATACACTAATTAAGGAACTCAGAAAGCAATACCCATTTAGTAAGATG 6664  
Qy 3351 GAYAMCTGAAGYMBRAAGTGCTTTCCAGGCCCTTAAAGAGSCCTTAAACCCCAAGYCCA 3410  
Db 6665 GAACACTGAAGTAGAAGTGCTTTCCAGGCC-----TAAACCCAGCCCA 6710  
Qy 3411 GTGTTAAGYTTGCCAACGGGCAAGACTTTTSTTVATRYRTCAAGAAAAAACHAGRAAY 3470  
Db 6711 GTGTTAAGYTTGCCAACGGGCAAGACTTTTCTTCATATGTACAG-AAAAAACHAGGAAT 6769  
Qy 3471 AGCTCTRGAGTCTTTACAGRTCCRAGGAYAGCTTGGAAACCYRTGGGCHYACTGAS 3530  
Db 6770 AGCTCTAGGAGTCTTTACAGATCCGAGGATGAGCTTGGCAACCTGTGGCGTACCTGAC 6829  
Qy 3531 TAAAGAAATGATGTAGTGGAAGGTTGRCYCTCATTTGTTTAYGGGTAGTGGTGGCAGT 3590  
Db 6830 TAAAGAAATTTGATGTAGTGGAAGGTTTGA CCTCATTTGTACGGGTAGTGGTGGCAGT 6889  
Qy 3591 AGCAGYKTAGTATCTGAAGCAGTTAAATAATAACAGGAGAGATCTTACTGTGTGGAC 3650  
Db 6890 AGCAGTCTTAGTATCTGAAGCAGTTAAATAATAACAGGAGAGATCTTACTGTGTGGAC 6949  
Qy 3651 ATCTCATGAGTGAAAYRGATACTCATCTGCTAAAGAGACTTGTGGCTGTACAGAACYVG 3710  
Db 6950 ATCTCATGATGTGAATGGCATACTCATCTGCTAAAGAGAGACTTGTGGCTGTACAGAAC 7009  
Qy 3711 TTTACTTAATTCAGGCTCTATTACTTGAAGGCCAGTGTGCRACCTGTGCACCTTTGTC 3770  
Db 7010 TTTACTTAATTCAGGCTCTATTACTTGAAGGCCAGTGTGCGACCTGTGCACCTTTGTC 7069  
Qy 3771 AACTCTTAAACCCAGYCNCAATTTCTTCCAGACATGAAGAAAAAGATARAAYATACTGTCA 3830  
Db 7070 AACTCTTAAACCCAGCAGATTTCTTCCAGACATGAAGAAAAAGATARAAYATACTGTCA 7129  
Qy 3831 ACAARTAAATTTCTAAACCTATGCCACTCGAGGGACCTTYPYTAGAGTTTCCYTTGACTGA 3890  
Db 7130 ACAAGTAATTTCTAAACCTATGCCACTCGAGGGACCTTTTAGAGGTTTCTTTGACTGA 7189  
Qy 3891 TCCYGACCTTCAACTTGTATCTGATGGAAGTTCTTTTGTAGAAAAAGGACTTCCGAAAG 3950

Db 7190 TCCCGACC-TCAACTTGTATACTGATGGAAGTCTCTTTGTAGAAAAAGGACTTCGAAAAG 7248  
Qy 3951 YGGGGTATGAGTGGTCACTGATTAATGGAATAYTTGAAAGTAATCCCTCACTCCAGGA 4010  
Db 7249 TGGGGTATGCACTGGTCACTGATTAATGGAATAYTTGAAAGTAATCCCTCACTCCAGGA 7308  
Qy 4011 CTAGTGTCTYAGCTTGCAGAACTAATAGCCYCTAYTKGGGCACTAGAAATTAGGAGAGRAA 4070  
Db 7309 CTAGTGTCTYAGCTTGCAGAACTAATAGCCCTCACTTGGGCACTAGAAATTAGGAGAGRAA 7368  
Qy 4071 AAAGGAAATATATATATACAGACTCTTARTATGCTYACCTAGTCTNTCCATGCCCATMRG 4130  
Db 7369 AAAGGCAATATA-ATACAGACTCTAATATGCTTACTTAGTCTCTCCATGCCCATGCG 7427  
Qy 4131 CAATATGSAAGAAAGGGAATTTCTTAATCTTCYAGGAGAACCTTATCMACATCAGGAAG 4190  
Db 7428 CAATATGSAAGAAAGGGAATTTCTTAATCTTCYAGGAGAACCTTATCMACATCAGGAAG 7487  
Qy 4191 CCATTTAGGABATTAATTAATGCGWGTACAGAACTTARAGAGTGTGMACTCTTACACTGY 4250  
Db 7488 CCATTTAGGAAATTAATTTTGGCTGTACAGAACTTARAGAGTGTGMACTCTTACACTGCC 7547  
Qy 4251 GGGGTCTATCANAAAGGAAAGGGAATASAAAGRAATYTGCCAAAGCAKATATTTGAAG 4310  
Db 7548 GGGGTCTATCAGAGGAAAGGGAATTAAGAGAACTGCAAGCATATATTTGAAG 7607  
Qy 4311 CMAAAGAGCTGCAAGGAGGACCTCCATTTAGAAATGCTTATTAATCTTCCCTTAGTAT 4370  
Db 7608 CCAAAAGAGCTGCAAGGAGGACCTCCATTTAGAAATGCTTATTAATCTTCCCTTAGTAT 7667  
Qy 4371 AGGGTAATCTCCGGAACCAAGCCCAAGTACTCAGCAGGAGAAACAGAAATCGGGAA 4430  
Db 7668 AGGGTAATCTCCGGAACCAAGCCCAAGTACTCAGCAGGAGAAACAGAAATCGGGAA 7727  
Qy 4431 CCTCAGAGGAGCTTTCTCCCTCGGACCGGTTAGCCACTCAAGAAAGGAAAAATATCTTT 4490  
Db 7728 CCTCAGAGGAGCTTTCTCCCTCGGACCGGTTAGCCACTCAAGAAAGGAAAAATATCTTT 7787  
Qy 4491 TGCTTGCAACTATCCAATGGAAATTA CTTAAAA CCTTTCATCAAAACCTTTTCACTTAGGA 4550  
Db 7788 TGCTTGCAACTATCCAATGGAAATTA CTTAAAA CCTTTCATCAAAACCTTTTCACTTAGGA 7847  
Qy 4551 TCGATAGCACCATTCARATGGCCAAATCATTTATTTACTTGGACCGCCCTTTTCAAAACTA 4610  
Db 7848 TCGATAGCACCATTCAGATGGCCAAATCATTTATTTACTTGGACCGCCCTTTTCAAAACTA 7907  
Qy 4611 TCAAGCARATATKTCAGGCGCTGTGAATGTGCAARAAAAATAATCCCTGCTCTATTCGCC 4670  
Db 7908 TCNAGCAGATAGTCAGGCGCTGTGAATGTGCAAGAGAAATAATCCCTGCTCTATTCGCC 7967  
Qy 4671 AAGCTCTTTCAGGAPAAACAAAACAGGCCATTTACCTTGRAPAAACCTTGGCAACTGATTT 4730  
Db 7968 AAGCTCTTTCAGGAGAACAAAGAACAGGCCATTTACCTTGGAGAGACTTGGCAACTGATTT 8027  
Qy 4731 TACCCAGAGCCCAACCTCAGGGATTTTCACTATCTACTGTCTGGGTARATACTTTCAC 4790  
Db 8028 TACCCAGAGCCCAACCTCAGGGATTTTCACTATCTACTGTCTGGGTARATACTTTCAC 8087  
Qy 4791 GGGTTGGCAGAGGCGCTTCCCTCTGAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACT 4850  
Db 8088 GGGTTGGCAGAGGCGCTTCCCTCTGAGGACAGAAAGGCCCAAGAGGTAAATAAGGCACT 8147  
Qy 4851 AGTTTCATGAAATAATTTCCAGATTCCGACTTCCCGAGGCTTACAGAGTGAACAATAGCCC 4910  
Db 8148 AGTTTCATGAAATAATTTCCAGATTCCGACTTCCCGAGGCTTACAGAGTGAACAATAGCCC 8207  
Qy 4911 TGCTTTCCAGGCCACAGTAAACCCAGGAGTATCCAGGGTTAGGTATACGATATCACTT 4970  
Db 8208 TGCTTTCCAGGCCACAGTAAACCCAGGAGTATCCAGGGTTAGGTATACGATATCACTT 8267  
Qy 4971 ACACCTGCGCTGAAGGCCACAGTCTCTCAGGGAAGGTTCGAGAAAAATGAATGAAYACTCAA 5030  
Db 8268 ACACCTGCGCTGAAGGCCACAGTCTCTCAGGGAAGGTTCGAGAAAAATGAATGAAYACTCAA 8327



Qy	5031	AGGACATCTAAAAAGCAAAACCCACAGGAAACCCACCTCACATGGCGCTGYTCTGTGGCCTAT	5090
Db	8328		
		AGGACATCTAAAAAGCAAAACCCACAGGAAACCCACCTCACATGGCGCTGTCTGTGGCCTAT	8387
Qy	5091	AGCCTTAAAAAGAAATCTGCAACTTTTCCCAAAAGACAGGACTTAGCCCATACGAAATGCT	5150
Db	8388		
		AGCCTTAAAAAGAAATCTGCAACTTTTCCCAAAAGACAGGACTTAGCCCATACGAAATGCT	8447
Qy	5151	GTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGGACCAAGACAGCCAACTAGTAT	5210
Db	8448		
		GTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGGACCAAGACAGCCAACTAGTAT	8507
Qy	5211	GCAGACATCACCTCTTAGGCCAAATATCAACAAGTCTTTAAAAACATTTACAAGGAACCTAT	5270
Db	8508		
		GCAGACATCACCTCTTAGGCCAAATATCAACAAGTCTTTAAAAACATTTACAAGGAACCTAT	8567
Qy	5271	CCCTGAGAAGGGAAAGAAAGAACTATTCCACCCWGTGACATGTTATTAGTCAAGTCCCTT	5330
Db	8568		
		CCCTGAGAAGGG--AAAAGAATATTCCACCCAAGTGACATGGTATTATTAGTCAAGTCCCTT	8625
Qy	5331	CYCTCTAAATTCCCCATCCCTAGATACATCTCTGGGAAGGACCCCTACCCAGATCATTTTATYT	5390
Db	8626	:   :	:   :
		CCCTCTAAATTCCCCATCCCTAGATACATCTCTGGGAAGGACCCCTACCCAGATCATTTTATCT	8685
Qy	5391	ACCCCAACTGCGGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTCAGTCAAA	5450
Db	8686		
		ACCCCAACTGCGGTAAAGTGGCTGGAGTGGAGTCTTGGATACATCACACTTCAGTCAAA	8745
Qy	5451	TCCTTGGATACGTGCCAAAGAACCTGAAAAATCCAGGAGACAAACGCTAGCTATTCTCTGTGAA	5510
Db	8746		
		TCCTTGGATACGTGCCAAAGAACCTGAAAAATCCAGGAGACAAACGCTAGCTATTCTCTGTGAA	8805
Qy	5511	CCTCTAGAGATTGGCCCTGCTCTTCAAAACAACCAAGGAGGAAA	5557
Db	8806	CCTCTAGAGATTGGCCCTGCTCTTCAAAACAACCAAGGAGGAAA	8852

## RESULT 8

RES001.6	
AAS84210	
ID	AAS84210 standard; cDNA; 6394 BP.
XX	
AC	AAS84210;
XX	
DT	13-FEB-2002 (first entry)
XX	
DE	DNA encoding novel human diagnostic protein #20014.
XX	
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.
OS	Homo sapiens.
XX	
PN	WO200175067-A2.
XX	
PD	11-OCT-2001.
XX	
PF	30-MAR-2001; 2001WO-US008631.
XX	
PR	31-MAR-2000; 2000US-00540217.
PR	23-AUG-2000; 2000US-00649167.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Dmanac RT, Liu C, Tang YT;
XX	
DR	WPI; 2001-639362/73.
DR	P-PSDB; ABG20023.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity.

Db 2559 GCAACCCAGAACGGACATCCAAAGTGATGGAAACGTTTCCCGCAAGACAAAAACGCCCC 2618  
Qy 641 TAAGACGTATTCTGABAAATTTGGGAMCAATTTGACCTTCAGACACTAA-GAAAGAAACGA 699  
Db 2619 TAAGACGTATTCTGGAGAAATTTGGGACCAATTTGACCTTCAGACACTAAAGGAAAGAAACGA 2678  
Qy 700 CTTATATTCTTCTGCAGTGGCCCTCGGCACCTCTGAGGGAAGTATAAAATTTATAACACCAT 759  
Db 2679 CTTATATTCTTCTGCAGTGGCCCTCGGCACCTCTGAGGGAAGTATAAAATTTATAACACCAT 2738  
Qy 760 CTTACAGCTAGACATCTCTTTTGTAGAA---AAGGCAAAATGGAGTGAAGTGCATTAAGTACA 816  
Db 2739 CTTACAGCTAGACATCTCTTTTGTAGAAAGAGGTTAAATGGAGTGAAGTGTATATGTACA 2798  
Qy 817 AACTTTCTTTTCAATTAAGAGACAACTCACAATTA-TGTAAAAAGTGTGATTTATGCCCC-- 873  
Db 2799 AACTTTCTTTTCAATTAAGAGACAACTCACAATTA-TGTAAAAAGTGTGATTTATGCCCCCT 2858  
Qy 874 TACAGGAAGCCTTCAGAGTCTACCTCCCTATCCG---AGCATCCCGACCTCTTCCCCA 929  
Db 2859 ACAGGAAGCCTTCAGAGTCTACCTCCCTATCCCGACATCCCCCAACTCTTCTCCA 2918  
Qy 930 MYTAAATAGGACCCCTTCAACCCAAATGGTCCAAAAGAGAGATAGACAAAAGGGTAAAC 989  
Db 2919 ACTATTAAGGACCCCTTCAACCCAAATGGTCCAAAAGAGAGATAGACAAAAGGGTAAAC 2978  
Qy 990 AGTGAACCAAGAGTGCANATATTCGCCAATATATGACCCCTCCCAAGCAGTGGGAGGAAG 1049  
Db 2979 AGTGAACCAAGAGTGCANATATTCGCCAATATATGACCCCTCCCAAGCAGTGGGAGGAAG 3038  
Qy 1050 AGAATTCGGCCAGCAGAGTGCATGTCYTTTTTCTCCAGACTTTAAAGCAATATAA 1109  
Db 3039 AGAATTCGGCCAGCAGAGTGCATGTCYTTTTTCTCCAGACTTTAAAGCAATATAA 3098  
Qy 1110 AACAGACTTAGGTAAATTTCTCAGATAAATCTGATGGCTATATGRTGTTTTACAGGGTT 1169  
Db 3099 AACAGACTTAGGTAAATTTCTCAGATAAATCTGATGGCTATATGRTGTTTTACAGGGTT 3158  
Qy 1170 AGGACAAATCTTTGATCTGACATGGAGAGATATATGTCATCTGCTTAATCAGACACTAA 1229  
Db 3159 AGGACAAATCTTTGATCTGACATGGAGAGATATATGTCATCTGCTTAATCAGACACTAA 3218  
Qy 1230 CCCCAATGAGAGAGTCCACCATTAATGACAGCTGAGRTTTCGGCATCTCTGGTATC 1289  
Db 3219 CCCCAATGAGAGAGTCCACCATTAATGACAGCTGAGRTTTCGGCATCTCTGGTATC 3278  
Qy 1290 TCAGTTCAGTCAATGGATANGATGACACAGAGAAAGANAATGATTTCCCAACAGGCC 1349  
Db 3279 TCAGTTCAGTCAATGGATANGATGACACAGAGAAAGANAATGATTTCCCAACAGGCC 3336  
Qy 1350 AGCARGCAGTTCCTAGTCTASACCTCAATTTGGGACACAGAAATCAGTAACATGGAGAT 1409  
Db 3337 AGCARGCAGTTCCTAGTCTAGACCTCATTTGGGACACAGAAATCAGTAACATGGAGAT 3396  
Qy 1410 TGGTCTGCAGACATTTGCTAATCTGTGTCTASAGGACTAAGGAAACTASGAGAAA 1469  
Db 3397 TGGTCTGCAGACATTTGCTAATCTGTGTCTAAGGACTAAGGAAACTAAGGAAACTAAGGAAA 3456  
Qy 1470 RTCTAYGAATTAATCAATGATGTCCACCAATACACAGGGGAAGGAAAGAAATCTCTACTG 1529  
Db 3457 ATCTAGGAATTAATCAATGATGTCCACCAATACACAGGGGAAGGAAAGAAATCTCTACTG 3516  
Qy 1530 CTTTCTCGAGAGACTAAGGGAGGCAATGAGGAAGCGTCTCTCTGTCACTGACTGACTCTTT 1589  
Db 3517 CTTTCTCGAGAGACTAAGGGAGGCAATGAGGAAGCGTCTCTCTGTCACTGACTGACTCTTT 3576  
Qy 1590 CTGAGGGCCAACTAATCTTAAGCGTAAAGTTTATCACTCAGTCAGCTGCAGACATTTAGAA 1649  
Db 3577 CTGAGGGCCAACTAATCTTAAGCGTAAAGTTTATCACTCAGTCAGCTGCAGACATTTAGAA 3636  
Qy 1650 AAAAATTTCAAAAGTCTCGCTAGGGCCGGAGCAAAAATCTTAGAAACCTTATGAACTTGGC 1709  
Db 3637 AAAAATTTCAAAAGTCTCGCTAGGGCCGGAGCAAAAATCTTAGAAACCTTATGAACTTGGC 3696

Qy 1710 AACYTCCGTTTTTTATATAGAGATCAGGAGGACGCGGAAACAGGACAAACCGGATTATTA 1769  
Db 3697 AACYTCCGTTTTTTATATAGAGATCAGGAGGACGCGGAAACAGGACAAACCGGATTATTA 3756  
Qy 1770 AAAAAAGGCGCACCGCTTTTAGTCAATGACCCCTCAGGCAAGTGGACTTTTGGAGGCTCTCGAA 1829  
Db 3757 AAAAAAGGCGCACCGCTTTTAGTCAATGACCCCTCAGGCAAGTGGACTTTTGGAGGCTCTCGAA 3816  
Qy 1830 AAGGAAAGCTGGGCAAAATTTGAATGCTAATAGGGCTTGTCTCAGTGCGGTCTTACAAG 1889  
Db 3817 AAGGAAAGCTGGGCAAAATTTGAATGCTAATAGGGCTTGTCTCAGTGCGGTCTTACAAG 3876  
Qy 1890 GACACTTTAAAAAGATTTGTCCAAAGTAGAGTAAAGCCGCTTCTGTCATGCCCCCTTAT 1949  
Db 3877 GACACTTTAAAAAGATTTGTCCAAAGTAGAGTAAAGCCGCTTCTGTCATGCCCCCTTAT 3936  
Qy 1950 TTCAAGGGAATCACTGGAAAGCCCACTGCCCCAGGGGACAAAGGTCTTTTTCAGTCAAG 2009  
Db 3937 TTCAAGGGAATCACTGGAAAGCCCACTGCCCCAGGGGACAAAGGTCTTTTTCAGTCAAG 3996  
Qy 2010 CCACTAACAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGGAAGCGCATCCCATG 2069  
Db 3997 CCACTAACAGATGATCCAGCAGCAGGACTGAGGGTGCCTGGGGGAAGCGCATCCCATG 4056  
Qy 2070 CCATCACCTCACAGAGCCCTGGGTATGCTTGAACATTTAGGGCCAGGAAGGTGTCTCC 2129  
Db 4057 CCATCACCTCACAGAGCCCTGGGTATGCTTGAACATTTAGGGCCAGGAAGGTGTCTCC 4116  
Qy 2130 TGGACACTGGTGGCTCTTCTTAGTCTTCTTCTGTCCTCCGGACAACTGTCCTCCAGAT 2189  
Db 4117 TGGACACTGGTGGCTCTTCTTAGTCTTCTTCTGTCCTCCGGACAACTGTCCTCCAGAT 4176  
Qy 2190 CTGTCACTATTCTCAGGGGGTCCNTAAGACGGGCAGTCACTAGATACATTTTTCAGGCA 2249  
Db 4177 CTGTCACTATTCTCAGGGGGTCC-TAAGACGGGCAGTCACTAGATACATTTTTCAGGCA 4235  
Qy 2250 CTAAGTATGAACTGGGAGCTTTATTTCTTTTACATGCTTTTCTTAATATGCTTGAAG 2309  
Db 4236 CTAAGTATGAACTGGGAGCTTTATTTCTTTTACATGCTTTTCTTAATATGCTTGAAG 4295  
Qy 2310 CCCCACTACCTTTTAGGGAGAGACATTTAGCAAAAGCAGGGGCCATTTATACACTGAA 2369  
Db 4296 CCCCACTACCTTTTAGGGAGAGACATTTAGCAAAAGCAGGGGCCATTTATACACTGAA 4355  
Qy 2370 CATAGGAAGGAAACACCCGTTTGTGTCCTTGTGTCCTTTGAGGAAAGAAATTAATCTGAAG 2429  
Db 4356 CATAGGAAGGAAACACCCGTTTGTGTCCTTGTGTCCTTTTAAATATGCTTGAAG 4414  
Qy 2430 TCTGGGCAACAGAGGACAAATATGAGCGAACAAAGATGCGCTCTGTTTCAGTTAAA 2489  
Db 4415 TCTGGGCAACAGAGGACAAATATGAGCGAACAAAGATGCGCTCTGTTTCAGTTAAA 4474  
Qy 2490 CTAAGGATTCACCTTCTCTTCCCTTACCAAGGAGTACCCCTCAGACCCCAAGGCCCAA 2549  
Db 4475 CTAAGGATTCCTGCTCTCTTCCCTTACCAAGGACATACCCCTTAGACCTGAGGCCCAA 4534  
Qy 2550 CAA--GGATTCAAAAGATTTGTAAGGACTTTAAAGGCCAAAGGCTTAGTAAACCATGCA 2607  
Db 4535 CAAAGGCTTCAAAAGATTTGTAAGGACTTTAAAGGCCAAAGGCTTAGTAAACCATGCA 4594  
Qy 2608 TAACTCCCTGAGTAATTCCTGATGGAATGAGGAGGACAGAAACCCAGTGGACAGTGG 2667  
Db 4595 TAACTCCCTGAGTAATTCCTGATGGAATGAGGAGGACAGAAACCCAGTGGACAGTGG 4654  
Qy 2668 AGGGTTAGTCAAGATCTCAGGATTTCAATGGAGGCGCTTGTCTCTTTTATACCCAGCTG 2727  
Db 4655 AGGGTTAGTCAAGATCTCAGGATTTCAATGGAGGCGCTTGTCTCTTTTATACCCAGCTG 4714  
Qy 2728 TACCTAGCCCTTATACTGTGMYTTCCCAAAATACAGAGAAAGCAGAGTGGTTTACSTCC 2787  
Db 4715 TACCTAGCCCTTATACTGTGCTTCCCAAAATACAGAGAAAGCAGAGTGGTTTACACTCC 4774

2788 TGACCTTTAGGATGCTCTTCTGTCATCCCTGTACATCCTGACTCTCAATCTCTGTTG 2847  
 4775 TGGACCTTAAGGATGCTCTTCTGTCATCCCTGTACATCCTGACTCTCAATCTCTGTTG 4834  
 2848 CCTTTGAAGATCTTCAAAACCCARCTCTCAACTCAGCTGACCTCTTATACCCCAAGGGT 2907  
 4835 CCTTTGAAGATCTTCAAAACCCARCTCTCAACTCAGCTGACCTCTTATACCCCAAGGGT 4894  
 2908 TCAGGATAGYCCCATCTATTTGGCCAGGATTTAGCCCAAGACTTGAGYCATYMTFAT 2967  
 4895 TCAGGATAGYCCCATCTATTTGGCCAGGATTTAGCCCAAGACTTGAGYCATYMTFAT 4954  
 2968 ACCTGGACACTCTTGTCTCTCGTAKGTGATG-ATTTACTTTTTRGCGCYCTTCAGAA 3026  
 4955 ACCTGGACACTCTTGTCTCTCGTAKGTGATG-ATTTACTTTTTRGCGCYCTTCAGAA 5014  
 3027 ACCTTGTGCATCAAGCCACCAAGCRCTCTTMAATTTCTCGCYACTGTGCTTACAWG 3086  
 5015 ACCTTATGCCATCAAGCCACCAAAATGCTCTTAAATTTCTTGTCTACCTTGTGCTACAAG 5074  
 3087 GTTTCCAAACGACCTGCTCTGCTCAAGGATTTAAATCTTAACTTAACTTAACTTAA 3146  
 5075 GTTTTCAAAACCAAGGCTGAGCTCTGCTCAAGGATTTAAATCTTAACTTAACTTAA 5134  
 3147 TCCAAAGGACCAAGGCTTCAAGGATTTAAATCTTAACTTAACTTAACTTAACTTAA 3206  
 5135 TCCAAAGGATCAAGGCTTCAAGGATTTAAATCTTAACTTAACTTAACTTAACTTAA 5194  
 3207 CYCAAAACCTTAAGCAACTAAGGATTTCTTGGCTTAAAGGATTTCTGCGCAAAATGG 3266  
 5195 CCCAAACCTTAAGCAACTAAGGATTTCTTGGCTTAAAGGATTTCTGCGCAAAATGG 5254  
 3267 ATTCCCAAGGATTTGCGCAAACTAAGGATTTCTTGGCTTAAAGGATTTCTGCGCAAA 3326  
 5255 ATTAA-CAGGTATGGCAAACTAAGGATTTCTTGGCTTAAAGGATTTCTGCGCAAA 5313  
 3327 AGCCATATCCATTTTARTAGATGGAATGGAATGGAATGGAATGGAATGGAATGGA 3386  
 5314 AGCCATATCCATTTTARTAGATGGAATGGAATGGAATGGAATGGAATGGAATGGA 5369  
 3387 AGAAGGCTTTAAACCAAGYCCAGTGTAAAGYTTGCCAAGGCAAGACTTTTGTVA 3446  
 5370 -----TAACCCAGCCGAGTGTAAAGTGTGCCAAGGCAAGACTTTTGTVA 5419  
 3447 TAYRTCAAGAAACCAAGYAGCTCTTGGAGTCTTACACAGTCCAGGAGYAG 3506  
 5420 TATGTACAGAA-AAAAGGAATAGCTCTAGGAGTCTTACACAGTCCAGGAGYAG 5478  
 3507 CTTCGAACCTGTGGCTGCTGAGTAAAGGATTTGATGTAGTGGCAAGGTTGTCYCA 3566  
 5479 CTTCGAACCTGTGGCTGCTGAGTAAAGGATTTGATGTAGTGGCAAGGTTGTCYCA 5538  
 3567 TTGTTTAYGGTGTAGTGGCTGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 3626  
 5539 TTGTTTAYGGTGTAGTGGCTGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 5598  
 3627 GGGAGAGATCTTACTGTGTGAGCATCTCATGAKGTGAAYRGATCACTCACTGCTAAAG 3686  
 5599 GGAAGAGATCTTACTGTGTGAGCATCTCATGATGTGAACAGATCACTCACTGCTAAAG 5658  
 3687 AGACTTGTGGCTGTGAGCAACYGTGTACTTAAATTCAGGCTCTATTAATTGAAGGCC 3746  
 5659 AGACTTGTGGCTGTGAGCAACYGTGTACTTAAATTCAGGCTCTATTAATTGAAGGCC 5718  
 3747 AGTGTGCRACCTGTGACCTGTGACCTTAAACCCAGYCNCAATTTCTTCCAGACATCA 3806  
 5719 AGTGTGCRACCTGTGACCTGTGACCTTAAACCCAGYCNCAATTTCTTCCAGACATCA 5778  
 3807 AGAAAGATARAAYATACTGTCAACAARTTAATTTCTCAAACTATGCCACTCGAGGGA 3866  
 5779 AAAAAGATAGACGTAATCTGTCAACAGGATTTGCTCAAACTATGCCCTCGAGGAAA 5838  
 3867 CCTTGTAGAGTTCCYTTGACTGATCCYAGACCTTCACTTGTATATCTGATGGAAGTCT 3926

5839 CCTTCTAGAGTTTCTTGTGATCCTGACC-TCAACTTGTATATTGATGGAAGTTCCT 5897  
 3927 TTGTAGAAAAGGACTTCGAAAAGYGGGTATGACAGTGGTCACTGATATAATGGAATAATTG 3986  
 5898 TTGTAGAAAAGGACTTCGAAAAGYGGGTATGACAGTGGTCACTGATATAATGGAATAATTG 5957  
 3987 AAAGTAATCCCTCACTTCCAGGAACCTAGTGTAGTGTGAGTGTGAGTGTGAGTGTGAGT 4046  
 5958 AAAATAATCCCTCACTTCCAGGAACCTAGTGTAGTGTGAGTGTGAGTGTGAGTGTGAGT 6017  
 4047 GGGCACTAGTAATTTAGGAGAGGAAAAGGGAATATATATACAGACTCTTAAGTATGCTT 4106  
 6018 CGGCACCTAAATTTAGAAAAGGGAATATATATATATATACAGACTCTTAAGTATGCTT 6077  
 4107 ACCTAGTCTTCCATGCCCATGMRGCAATATGSAAGAAAGGGAATTTCTTAACCTTCYAGR 4166  
 6078 ACCTAGTCTTCCATGCCCATGMRGCAATATGSAAGAAAGGGAATTTCTTAACCTTCYAGR 6137  
 4167 GAACACCTATCAACATCAGGAAGCATTAGGAGATTTATTTATGTCGCTGTACAGAACTTA 4226  
 6138 GAACATCTTCAACATCAGGAAGCATTAGGAGATTTATTTATGTCGCTGTACAGAACTTA 6197  
 4227 RAGAGTGTGAGTCTTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4286  
 6198 RAGAGTGTGAGTCTTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6257  
 4287 RGAAYTGCCCAAGCAKATATTGAAGCAAAAGAGCTGCAAGGAGGAGGAGGAGGAGGAGG 4346  
 6258 GGAACCCGACAGGATATTGAAGCAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6317  
 4347 TGCTTATTAACTTCCCTTAGTATAGGGTAATCCCTT 4383  
 6318 TGCTTATTAACTTCCCTTAGTATAGGGTAATCCCTT 6354

RESULT 9  
 AAS76475  
 ID AAS76475 standard; cDNA; 4349 BP.  
 XX AC AAS76475;  
 XX AC AAS76475;  
 XX DT 13-FEB-2002 (first entry)  
 XX DNA encoding novel human diagnostic protein #12279.  
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX OS Homo sapiens.  
 XX PN WO200175067-A2.  
 XX PD 11-OCT-2001.  
 XX PF 30-MAR-2001; 2001WO-US008631.  
 XX PR 31-MAR-2000; 2000US-00540217.  
 XX PR 23-AUG-2000; 2000US-00649167.  
 XX (HYSE-) HYSEQ INC.  
 XX Drmanac RT, Liu C, Tang YT;  
 XX WPI; 2001-639362/73.  
 XX P-PSDB; ABG12288.  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX Claim 1; SEQ ID NO 12279; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 4349 BP; 1319 A; 1061 C; 941 G; 1026 T; 0 U; 2 Other;

Query Match 45.8%; Score 3431; DB 5; Length 4349;  
 Best Local Similarity 87.3%; Pred. No. 0;  
 Matches 3974; Conservative 148; Mismatches 195; Indels 237; Gaps 28;

QY 115 ACTGCACCTCTCTGCTCCATGTTCTTAGCGCTCGAGCTGAGCTTTTGTCTACCGTCCAC 174  
 DB 1 ACTGCACCTCTCTGCTCCATGTTCTTAGCGCTCGAGCTGAGCTTTTGTCTACCGTCCAC 60

QY 175 CACTGCTGTTTGCCACCACCGCANACCTGCCGCTGACTCCCATCCCTCTGGAATCTCGAG 234  
 DB 61 CACTGCTGTTTGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGAATCTCGAG 120

QY 235 GGTGTCGCTGTGCTCTGATCCAGGA---RGCGCCCATTTGCCGCTCCCAATGGGCTA 291  
 DB 121 GGTGTCGCTGTGCTCTGATCCAGCGAGCGCGGCCCATTTGCCACTCCCAAGTCGGGCTA 180

QY 292 AAGGCTTGCCATTTGTCCTGAC--GGCTAAGTGCCT--GGTTTGTCTTAATGAGCTGAA 349  
 DB 181 GAGGCTCACCATTGTTCTGACAGCTTAAGTGCCTGGGTTTCATCTTAATCGAGCTGAA 240

QY 350 CACTANTCAGTGGGTTCATGGTTCTCTCTGTGACCCACCGCTTCTAATAKAACATAAA 409  
 DB 241 CACTAATCAGTGGGTTCATGGTTCTCTCTGTGACCCACCGCTTCTAATAKAACATAAA 300

QY 410 CACTTACCCATGGCCCAAGATTCATTCCTTGGAAATCGGTGAGSCAACGNACTCCAGG 469  
 DB 301 CACTTACCCATGGCCCAAGATTCATTCCTTGGAAATCGGTGAGSCAACGNACTCCAGG 359

QY 470 TCAGAGAATACGAGCTTGCCACCATTTGGAAGCGGCTGCTACRCTCTTGGAAAGTGT 529  
 DB 360 TCAGAGAATACGAGCTTGCCACCATTTGGAAGCGGCTGCTACRCTCTTGGAAAGTGT 419

QY 530 TCACCACCATCTTTGGAGCTCTGTGAGCAAGGACCCCGGTGACATTTTGGSCACCAMS 589  
 DB 420 TCACCACCATCTTTGGAGCTCTGTGAGCAAGGACCCCGGTGACATTTTGGSCACCAMS 479

QY 590 RACGACATCCMAAGTATGGGAAAGTTTCCCGGCAAGCAAAAACGCCCTTAAGACGTA 649  
 DB 480 RACGACATCCMAAGTATGGGAAAGTTTCCCGGCAAGCAAAAACGCCCTTAAGACGTA 539

QY 650 TTCTCGARAATTGGGACCAATTTGACCTCAGACACTAAGAAGAAACGACTTATATCT 709  
 DB 540 TTCTCGARAATTGGGACCAATTTGACCTCAGACACTAAGAAGAAACGACTTATATCT 599

QY 710 TCTGCACTGCCCGCTGGCACTCTCTGAGGGAAGTATAAATTAACACCATCTTACAGCTA 769  
 DB 600 TCTGCACTGCCCGCTGGCACTCTCTGAGGGAAGTATAAATTAACACCATCTTACAGCTA 659

QY 770 GACYTCTTTTGTAGAAAGGCAAAATGGAGTGAAGTGCATAGTACAACTTTCTTTTCA 829  
 DB 660 GACCTCTTTTGTAGAAAGGCAAAATGGAGTGAAGTGCATAGTACAACTTTCTTTTCA 719

QY 830 TTAAGAGCAACTCACAATTAATGTAAAAAGTGTGATTTATCCCTTACAGGAAGCTTCAAG 889  
 DB 720 TTAAGAGCAACTCACAATTAATGTAAAAAGTGTGATTTATGCCCTTACAGGAAGCTTCAAG 779

QY 890 AGTCTACTCTCCCTATCCAGCATCCCGATCCCTTCCCCMYTAATAGGACCCCTTC 949  
 DB 780 AGTCTACTCTCCCTATCCAGCATCCCGATCCCTTCCCCMYTAATAGGACCCCTTC 839

QY 950 AACCCTAATGTCCTAAAGGAGATAGACAAAGGCTAAACAGTGAACCAAGAGTGCCAA 1009  
 DB 840 AACCCTAATGTCCTAAAGGAGATAGACAAAGGCTAAACAGTGAACCAAGAGTGCCAA 899

QY 1010 TATTCCCAATTAATGACCCCTCCAGCAGTGGGAGGAGAAATTCGGCCCGCAGAG 1069  
 DB 900 TATTCCCAATTAATGACCCCT--CCAAGCAGTGGGAGGAGAAATTCGGCCCGCAGAG 958

QY 1070 TGCATGTCYTTTTTCTCCAGACTTAAGCAATAAACAACAGCTTAGTAAATTTCT 1129  
 DB 959 TGCATGTCYTTTTTCTCCAGACTTAAGCAATAAACAACAGCTTAGTAAATTTCT 1018

QY 1130 CAGATAATGTCGCTATATGCTGTTTTCAGGGTTAGGACAAATCTTTGATCTGA 1189  
 DB 1019 CAGATAATGTCGCTATATGCTGTTTTCAGGGTTAGGACAAATCTTTGATCTGA 1078

QY 1190 CATGAGAGATATATGCTACTGCTAAATCAGACACTAACCCCAAAATGAGAGAGTGCC 1249  
 DB 1079 CATGAGAGATATA--ATGCTACTGCTAAATCAGACACTAACCCCAAAATGAGAGAGTGCC 1137

QY 1250 ACCATACTGAGCCTGAGGTGTTGGCGATCTCTGATATCTCAGTCAGTCAATGGATAN 1309  
 DB 1138 ACCATACTGAGCCTGAGGTGTTGGCG--TCTCTGATATCTCAGTCAGTCAATGGATAN 1196

QY 1310 GGATGACAAAGAGGAAAGANAATGATTCGCCACAGGCCAGCAGCAGTTCGCCAGTCTA 1369  
 DB 1197 GGATGACAAAGAGGAAAGANAATGATTCGCCACAGGCCAGCAGCAGTTCGCCAGTCTA 1256

QY 1370 SACCCTCATTTGGGACACAGAAATCAGTAAACATGGGAGATTTGGTCTGCAGACATTTGCT 1429  
 DB 1257 GACCCTCATTTGGG--ACACAGAAATCAGAAACAT--GGAGATTTGGTCTGCAGACATTTGCT 1312

QY 1430 AACTTGTGCTGCTAAGGAACTAAGGAAACTASGAGAAARTCTATGAAATTAATCTGA 1489  
 DB 1313 AACTTGTGCTGCTAAGGAACTAAGGAAACTTAGGAAG--AGTCTATGAATTAATCTGA 1371

QY 1490 TGTCCACCATAAACAGGGGAGGAGAAATCTCTACTGCTTTCTGGAGAGACTAAGG 1549  
 DB 1372 TGTCCACCATAAACAG--GGGAGGAGGAGAAATCTCTACTGCTTTCTGGAGAGACTAAGG 1430

QY 1550 GAGGCATTTGAGGAAGCGTCTCTCTGCTCAGCTGATCTTCTGAGAGGCCAACTAATCTTA 1609  
 DB 1431 GAGGCATTTGAGGAAGCGTCTCTCTGCTCAGCTGATCTTCTGAGAGGCCAACTAATCTTA 1490

QY 1610 AAGGTAAAGTTTATCATCTCAGTCAGCTGACAGATTAAGAAAAAATCTCAAAAGTCTGCGG 1669  
 DB 1491 AAGGTAAAGTTTATCATCTCAGTCAGCTGACAGATTAAGAAAAAATCTCAAAAGTCTGCGG 1550

QY 1670 TAGGCCCGGAGCAAACTTAGAAACCTTATGAACTTCGGCACTTCGGTTTTTTTATAATA 1729  
 DB 1551 TAGGCCCGGAGCAAACTTAGAAACCTTATGAACTTCGGCACTTCGGTTTTTTTATAATA 1610

QY 1730 GAGATCAGGAGGAGCAGCGGAAACAGGACAAACCGGATTTAAAAAAGGCCCAACCGCTTTA 1789  
 DB 1611 GAGATCAGGAGGAGCAGCGGAAACAGGACAAACCGGATTTAAAAAAGGCCCAACCGCTTTA 1670

QY 1790 GTCATGACCTCAGGCAAGTGAATTTGGAGGCTCTGGAAAAAGGAAAAAGCTGGGCAAT 1849  
 DB 1671 GTCATGACCTCAGGCAAGTGAATTTGGAGGCTCTGGAAAAAGGAAAAAGCTGGGCAAT 1730

QY 1850 TGAATGCTAATAGGCTTGTCTTCCAGTGGCTCTTCAAGGACACTTTAAAAAAGATTTGT 1909

Db 1731 TGAATGCTTAATAGGCTTCTTCCAGTGGCGGTCTACAAGGACACTTTAAAAAGATTGT 1790  
Qy 1910 CCAAGTAGAAGTAGCGGCCCTTCGTCCATGCCCCCTTATTCAAGGGAATCACTGGAAG 1969  
Db 1791 CCAAGTAGAA----- 1800  
Qy 1970 GCCCACTGCCAGGGGACAAAGTGCTTTTGTAGTTCAGNAGCCACTAACAGATGATCCAG 2029  
Db 1801 ----- 1800  
Qy 2030 CAGCAGGACTGAGGGTGCCTGGGGCAAGCGCCATCCCATGCCATCCCTCACAGAGCCC 2089  
Db 1801 -----AGGCC 1805  
Qy 2090 TGGGTATGCTTGACCAATTGAGGGCGCAGGAAGGT-----TGTCTCTGGACACTGTGGCGGT 2146  
Db 1806 CAGGTATGCTTGACCAATTGAGAGCCAAAGAGGTAACTGTCTCCCTGGACGCGGTGGCGCC 1865  
Qy 2147 TTTCTTAGTCTTACTCTTCTGTCGCCGGAACAACCTGCTCCAGATCTGTCACTATTCTGAGG 2206  
Db 1866 TTTCTAGTCTTACTCTCTGTCGCCGGAACAACCTGCTCCAGATCTGTCACTATCC---GA 1922  
Qy 2207 GGSTCCNTAAGCGGCGAGTCACTAGATACCTTTTCCAGCCACTAAGTTTATGAATGGG 2266  
Db 1923 GGGGCTCTAGGACAGCGAGTCACTAGATAC-TTCTTCCAGCCACTAAGTTGTG-ACTGGG 1980  
Qy 2267 GAGCTTTATTTTTCATGCTTTTCTAATTATGCTTGAAGCCCCACCTACCTTGTAG 2326  
Db 1981 GAGCTTTTACCTTTTTCATGCTTTCTAATTATGCTTGAAGCCCTCACTCCCTTGTAG 2040  
Qy 2327 GGAGAGACATTTAGCAAGAGCGGGCCATTATACCTGACATAGGAGAGGAACAC 2386  
Db 2041 AGAGAGAGATTTCTAGTAAAGAGCGGGCCATTATACCTGAAACATAGGAGAGGAACAC 2100  
Qy 2387 CCGTTTGTGTGTCCTCTGCTTGAAGAGGAATTAATCTCGAAGTCTGGGCAACAGAGGA 2446  
Db 2101 CCAATTGTGCT-CTCTGTTTGAAGAGGAATTAGTCTTGAACTCTGGGCAACAGAGGA 2159  
Qy 2447 CAATATGGACGACCAAGAAATCCCGTCTGTTCAAGTTTAACTAAAGGATTTCCACTTC 2506  
Db 2160 CAATATGGCAAG-CAAGAAGTCCCATTTTGTCCAAGTTAATCAAGGATTTCCACCTC 2218  
Qy 2507 CTTTCCCTCAAGAGGAGTACCCCTCAGACCAAGGCCCAACAGAGATTCCAAAGAT 2566  
Db 2219 CTTTCCCTATCAAGGAGTACCCCTCAGACCGGAGGCCCAACAAAGATCTCCAAAGAT 2278  
Qy 2567 TGTAAAGGACTTAAAGCCCAAGGCTTAGTAAACCATGCAATACCTCCCTGCAATATTC 2626  
Db 2279 TGTTAAGGACCTAAAGCCCAAGGCTTAGTAAACCATGCAATACCTCCCTGCAATATTC 2338  
Qy 2627 CGTAGTGGATTGAGGAGGCACAGAAACCCAGTGGACAGTGGAGGGTTAGTGCAGATCTC 2686  
Db 2339 -----AATTTTAGGATACAGAAACCCAAAGACAGTGA-GGTTAGTGCAGATCTC 2390  
Qy 2687 AGGATTAATCAATGGAGGCCGTTGCTCTTTTATACCAGCTGTACCTAGCCCTTATCTGT 2746  
Db 2391 AGGATTAATCAAT-GAGGCTGTTTTCTCTATACCAGCTATACCTAGCCCTTATCTCT 2449  
Qy 2747 GWTTCCTCAATACAGAGAGCAGAGTGGTTTACATCTCGACTTTCAGATGAGTCCCTT 2806  
Db 2450 GCTTTCCCAATACAGAGAGCAGAGTGGTTTACAGTCTCGGATCTTAAGGATGCTT 2509  
Qy 2807 CTTCTGATCCCTGTATCATCTGACTCTCAATTTCTTTGTTGCTTTTGAAGATACTTCAAA 2866  
Db 2510 TTTCTGATCCCTGTATCATCTGACTCTCAATTTCTTTGTTGCAATTTGAAGATCTTCAAA 2569  
Qy 2867 CCCATCATCTCAACTCACTCGGACTTTTTTACCCCAAGGGTTCAAGGATAGYCCCCATCT 2926  
Db 2570 CCCAATCTCAACTCACTCGGACTTTTTTACCCCAAGGGTTCAAGGATAGTCCCCATCT 2629  
Qy 2927 ATTTGGCCAGGATTTAGCCCAAGACTTGAGYCAATTTTATCATCTGACACTCTTGTCTC 2986

Db 2630 ATTTGGCCAGGATTTAGCCCAAGACTTTGAGCCAGTTCTCATACCTGGACACTCTGTCTC 2689  
Qy 2987 TCRGTATGCTGGATGATTTATCTTTTTCGTCGCTTTCAGAAACCTTGTGCAATCAAGCCAC 3046  
Db 2690 TCAGTGCATGGATGATTTATCTTTTAGTCCCGTTTCAGAAACCTTGTGCAATCAAGCCAC 2749  
Qy 3047 CCAAGCRCTCTTMAATTTCTCGCYACCTGTGCTACAMGGTTTCCAAACSAARAGCTCA 3106  
Db 2750 CCAAGCGCTCTTAAATTTCTCACTACCTGTGCTAAAGGTTTCCAAACCAAGGCTCA 2809  
Qy 3107 RCTCTGCTCACAGCAGGTTAAATACTTAGGRCCTAARATTTATCCAAAGGACCAARGGCCCT 3166  
Db 2810 GCTCTGCTCACAGCAGGTTAAATGCTTAGGCTTAAATTTATCCAAAGTCAACAGGCCCT 2869  
Qy 3167 CAGTGAAGAAATATCCAGCCTATACCTGCTTATCTCATCTCYCAAAACCTTAAAGCAACT 3226  
Db 2870 CAGTGAAGAAACGTTATCCAGCCTATACCTGCTTATCTTATCCAAACCTTAAAGCAACT 2929  
Qy 3227 AAGGRRTTCTTGGCCTAAAYAGGTTCTGCCGAAATGATGATTTCCAGGTACGGCAAA 3286  
Db 2930 AAGAGGTTCTTGGCCTAACAGGTTTCTGCCGAAATGATGATTTCCAGGTACGGCAAA 2988  
Qy 3287 TAGCCAGGYCATTAATACATAAATAAGGAACTCAGAAAGCCAAATACCAATTTATATA 3346  
Db 2989 TAGCCAGACGTTATATACGCTTAATTAAGGAACTCAGAAAGCCAAATACCAATTTATATA 3048  
Qy 3347 GATGAYACTGAAYMRAAGTGGCTTTCCAGGCCCTTAAAGAGGCCCTTAAACCCAAAGY 3406  
Db 3049 GATGGACACTCAAGCAAGAGCAGCTTTCCATGCCCTTAAAGAGGCCCT--AATGCAAGC 3106  
Qy 3407 CCGAGTTTAAAGYTTGCCAAACGGGCAAGACTTTTCTCATATGTCACAAAAACAAACA 3454  
Db 3107 CCGAGTGTAGCTTGTCTAACGGGCAAGACTTTTCTCATATGTCACAAAAACAAACA 3166  
Qy 3455 -----AGAAAAAACAGRAAYAGCTCTGGAGTCTTACACAGRTCCRAGGAYGA 3505  
Db 3167 AACAAAGAAAAAACAATGATAGCTCTAAGAGTCTCTGGCTGGTCTGAGGACGA 3226  
Qy 3506 GCTTGCACCCYRTGGCRVACCTGASTAAGGAAAYTGATGTAGTGGCAAGGGTTGTCCTC 3565  
Db 3227 GTTTGCACCCCTGGCATACCTGA-TAAGGAAATGATGTAGTGGCAAGGGTCAAGCTC 3285  
Qy 3566 ATTTGTTTAYGGTGTAGTGGCAGTAGCAGTCTGATGTATCTGAAGCAGTTAAATATATAC 3525  
Db 3286 ATTTGTTTACGGTGTAGTGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTAAATATATAC 3345  
Qy 3626 AGGGRAGAGACTTCTAGTGTGCGACATCTCATGAGTGAAYRGCACTACCTGCTTAAG 3685  
Db 3346 AGGGAAGAAATCTTACTGTGTGACATCTCATGATGTGAATGGCATACTCACTGTCTAAG 3405  
Qy 3686 GAGACTTGTGGCTGTGACAGCAACCYGTTTACTTTAAATRTCAGGCTCTATTACTTGAARGGC 3745  
Db 3406 GAGACTTGTGGCTGTGACAGCAACCATTTACTTTAAATATCAGGCTCTATTACTTGAAGGC 3465  
Qy 3746 CAGTGTGCRACCTGTGCACTTGTGCAACTCTTTAACCCAGYCNACTTTCTTCAGACAAATG 3805  
Db 3466 CAATGTGCGACTGTGCACTTGTGCACTTCTTAACTGAGACATTTCTTCCACAAATG 3525  
Qy 3806 AAGAAAGATATAAATAACTGTCAACAAATTTCTCAAAACCTATGCACTCGAGGGG 3865  
Db 3526 AAGAAAGATATAAATAACTGTCAACAAATTTGCTCAAAACCTATGCACTCGAGGGG 3585  
Qy 3866 ACTTGTAGAGTTCCTTTGACTGTATCGAGCTTCAACTGTATATCTGATGCAAGTCTC 3925  
Db 3586 ACATTTAGAGGTTCCCTTGACTGTATCCAAACCT-ACCTTGTATCTGATGGAAGATCC 3644  
Qy 3926 TTTGTAGAAAAAGGACTTCGAAAAAGYGGGTTATGCAAGTGGTCACTGATTAATGGAATYTT 3985  
Db 3645 TTTGTAGAAAAAGGACTTCGAAAAAGTGGGTTATGCAAGTGTGATTAATGGAATYTT 3704  
Qy 3986 GAAAGTATCCCTCACTCCAGGAACTAGTGTCTGAGTGAAGAACTAATAGCCYCAAT 4045  
Db 3705 GAAAGTATCCCTCACTCCAGGAACTAGTGTCTGAGTGGCAAACT-----TCACT 3755



Qy	4046	KGGCACCTAGAAATTAGGAGAGGAAAAAGGYYAAATATATATACAGACTCTRTATGCT	4105
Db	3756	CGGCGCTAGAAATTAGGAGAGGAAAAGGGTAAATATATACAGACTCTAAGTATGCT	3815
Qy	4106	YACCTAGTCNTCCATGTCGCCATGMRGCGAATATGSRNAGAAAGGGAATTCCTAACTTCYCGAG	4165
Db	3816	TACCTAGTCCTCCATGTCGCCATGCGAGTAAATATGGAGAGAAAGGGAATTCCTAACTTCCTGAG	3875
Qy	4166	RGACACCTATCAMACATCAGGAGCCATTAGGARATATATATYGGGCGGTACAGAAACCT	4225
Db	3876	GGAAACCTATCAACATCAGGAAGCCATTAGGGAATTTATTTGGGCGGTACAGAAACCT	3935
Qy	4226	ARAGAGTGTGMAGTCTTACATCTGCGGGGTCAATCANAAAGGAAAGGAAAAATASAA	4285
Db	3936	AAAGAGTGTGGAGTCTTACATCTGCGGGGTCAATCAGAAAGGAAGGAATGAA	3995
Qy	4286	GRGAAATGCCAAGCAKATATTTGAAGCMAAAGAGCTCAAGCGAGGACCTCCATTAGAA	4345
Db	3996	AGBNACGCCAAGCAGATATTTGAAGCCAAAGAGCTGCAAGGCGAGGACCTCCATTAGAA	4055
Qy	4346	ATGCTTATTAAACTTCCCTTAGTATAGGTAATCCCTTTCGGGAAACCAAGCCCCAGTAC	4405
Db	4056	ATGCTTATAGAAGACCCCTAGTATGGGTAAATCCCTCTGGGAAACCAAGCCCCAGTAT	4115
Qy	4406	TCAGCAGGAGNAAACAGAAATGGGGAACCTCAGGAGG---CAGTTTCTCCCTCGGAGCG	4462
Db	4116	TCAGAAAGAAGAAATAGAAATGGGGAATCTCAGCAGGACATAGTTTCTCCCTCAGGATGG	4175
Qy	4463	TTAGCCCATGAAGAAGGGAAATACTTTTGCTGCAACTATCCAAATGGAAATTACTTTAAA	4522
Db	4176	CAAGCCACCGNACAAGGAAANATACCTTTTGCTCGAGCTAACCAATGGAAATTACTTTAAA	4235
Qy	4523	ACCTTTCATCAAACTTTTCACTTAGGCGATCGATAGCACCCATCARATGGCCAAATCATTA	4582
Db	4236	ACCACTCACBAAACTTTCACTTTATGCAATTGATAGCACCCATCAGATGGCCAAATCTGTA	4295
Qy	4583	TTTTACTGGACCAAGCCCTTTTCAAAACTATCAAGCARATAKTTCAGGGCCTGTGAA	4636
Db	4296	TTTTACTGGACCAAGCCTTTTCAAACTCTCAAGCAGATAGTTGGGCGCTGTAAA	4349

RESULT 10

AAS68626  
 ID AAS68626 standard; CDNA; 7466 BP.

AC AAS68626;

DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #4430.

xx  
kw Human; chromosome mapping; gene mapping; gene therapy; forensic;  
kw food supplement; medical imaging; diagnostic; genetic disorder; ss.  
kw Human; chromosome mapping; gene mapping; gene therapy; forensic;  
kw food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX  
PN  
WQ200175067-A2XX  
PD  
11-OCT-2001XX  
PF 30-MAR-2001. 2001WO-IIS008631

XX  
PP 31-MAR-2000: 2000US-00540217

PR 23-AUG-2000; 2000US-00649167.  
yy

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

PS Claim 1: SEO ID NO 4430: 103pp: English:

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AA364157-AAS94584 represent novel human diagnostic coding sequences of the invention. Notice: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published/pct/sequences](http://wipo.int/pub/published/pct/sequences)

Sequence 7466 BP: 2217 A; 1880 C; 1614 G; 1754 T; 0 U; 1 Other;

Query Match 44.2% Score 3309.6 DB 5 Length 7466

Best Local Similarity 95.7%; Pred. No. 0;  
Matches 3417. Conservative 98. Mismatches

Qy	1	CAACAATCGGGATATAAAACCCAGGCATTGAGCTGGCAACAGCAGCCCCCTTTGGTGCC	60
Db	3924	CAATGATCGGGATATAAAACCCAGGCATTGAGCTGGCAACAGCAGCCCCCTTTGGTGCC	3983
Qy	61	CTTCCCTTTGTATGGGAG--CTGTTTTCATGCTATTTCCTCTATTAAATCTTTGCAACTG	118
Db	3984	CTTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCCTCTATTAAATCTTTGCAACTG	4043

Qy	119	CAC	TCT	TC	GGT	CCAT	GT	TTCT	TAC	GGC	TC	GAG	CT	TTT	TG	CTC	A	CG	TC	CA	CC	ACT	178
Db	4044	CAC	TCT	TC	GGT	CCAT	GT	TTCT	TAC	GGC	TC	GAG	CT	TTT	TG	CTC	A	CG	TC	CA	CC	ACT	4103
Qy	179	GCT	TTT	TG	CC	AC	CG	CAN	AC	TG	CG	CT	GAC	T	CC	AT	C	CC	CT	CT	GAT	C	238
Db	4104	GCT	TTT	TG	CC	AC	CG	CAG	AC	TG	CG	CT	GAT	C	CC	AT	C	CC	CT	CT	GAT	C	4163
Qy	239	T	CG	CT	TG	TG	CT	CT	GAT	CC	AG	CG	G	CC	CA	T	T	G	CG	CT	CC	AA	298
Db	4164	T	CG	CT	TG	TG	CT	CT	GAT	CC	AG	CG	G	CC	CA	T	T	G	CG	CT	CC	AA	4223
Qy	299	G	CA	T	TG	TC	G	CA	CG	G	CT	T	A	G	T	T	C	T	A	A	T	T	358
Db	4224	G	CA	T	TG	TC	G	CA	CG	G	CT	T	A	G	T	T	C	T	A	A	T	T	4283
Qy	359	CT	G	G	T	T	CC	AT	G	T	T	CT	CT	G	A	C	CC	A	C	G	G	T	409
Db	4284	CT	G	G	T	T	CC	AT	G	T	T	CT	CT	G	A	C	CC	A	C	G	G	T	4343
Qy	410	C	A	T	T	A	C	C	A	T	T	CC	AT	T	CC	T	T	G	G	A	T	CC	469
Db	4344	C	A	T	T	A	C	C	A	T	T	CC	AT	T	CC	T	T	G	G	A	T	CC	4402
Qy	470	T	C	A	G	A	T	A	C	G	CT	T	G	C	A	C	A	T	T	G	A	G	529
Db	4403	T	C	A	G	A	T	A	C	G	CT	T	G	C	A	C	A	T	T	G	A	G	4462
Qy	530	T	C	A	C	C	A	C	A	T	T	G	G	A	G	CT	T	C	T	G	A	C	589





Qy	2750	TTCCCAATACACGAGGAAGCAGAGTGCGTTTACASTCCTCGACCTTWMAGANTGCTTCTT	2809
Db	6679	TTCCCAATACACGAGGAAGCAGAGTGCGTTTACATCTCCTCGACCTTAAAGANTGCTTCTT	6738
Qy	2810	CTGCATCCCTGTATACATCTCGACTCTCAATCTTCTGTTTGCCTTTGGAAGATACTTCAAAACC	2869
Db	6739	CTGCATCCCTGTATACATCTGACTCTCAATCTTGTGTTGCTTTGGAAGATACTTCAAAACC	6798
Qy	2870	ARCATCTCAACTCACCTGGACTRPTTTTACCCAAAGGGTTCAGGGATAGYCCCCATCTATT	2929
Db	6799	AACATCTCAACTCACCTGGACTGTTTTACCCAAAGGGTTCAGGGATAGCCCCCATCTATT	6858
Qy	2930	TGGCCAGGCATTAGCCCAAGACTTGAGYCAATYMTCATACCTGACACTCTTTGTCCTTCTC	2989
Db	6859	TGGCCAGGCATTAGCCCAAGACTTGAGCCCAATCCTCATCTGAGCA--CTTTGTCCTTTC	6916
Qy	2990	GTAKGTGGATGATTTACTTTTTRGCYGCYRTTCAGAAACCTTGTGCCATCAAGCCACCCA	3049
Db	6917	GTAGTGGATGATTTACTTTTGGCGGCCAATTGAGAACTTTGTGCCATCAAGCCACCCA	6976
Qy	3050	AGRCTCTTTMAATTTCTCGCYACCTGTGGCTACAWGGTTTCCAAACGAPARGCTCARCT	3109
Db	6977	AGOGCTCTTCAATTTCTCGCTACTCTGTGGCTACATGTTTCCAAACCAAGGCTCAACT	7036
Qy	3110	CTGCTCACAGCAGGTTAAATACTTAGRCCTAABATTTACAAAGGCACCAAGGCCCTCAG	3169
Db	7037	CTGCTCACAGCAGT----TACTTAGGGCTAAAAATTTACAAAGGCACCAAGGGCCCTCAG	7092
Qy	3170	TGAGGAARYATCCAGCCTACTACTGGCTTATCCTCATCYAAAAACCTTAAAGCAACTAAG	3229
Db	7093	TGAGGAACATCCAGCCTACTACTGGCTTATCCTCATCCCAAAACCTTAAAGCAACTAAG	7152
Qy	3230	RGRRTCTCTGGCCTAAYAGGYTCTGCCAAWATGGATTCGCCAGTWTGGCRAAATAG	3289
Db	7153	GGGATCTCTGGCGTAATAGGTTCTGCGCAAAATGGATT--CCCAGGTTTGGCGAAATAG	7211
Qy	3290	CCAGGYCATTAWATACASTAATTAAGAGAACTCAGAAAGCCAAATACCCTTTTARTAGAT	3349
Db	7212	CCAGGTCATTTAAATACACTAATTAAGGAACCTCAGAAGCCAAATACCCTTTTGTAGTAAAT	7271
Qy	3350	GGAYMCTGAAGYMRAGTGCTTTCCAGGCCCTTAAAGAGGGCTTTAAACCCCAAGYCCC	3409
Db	7272	GGACAACTGAAGTAGAAGTGCTTTTCAGGCC-----TAACCCCAAGCCCC	7317
Qy	3410	AGTGTTAAGYTTGCCAACRGGGCAAGACTTTTSTTYATAVETCACAGAAAAAACAGRAA	3469
Db	7318	AGTGTTAAGYTTGCCAACRGGGCAAGACTTTTCTTTCATATGTCTCAG-AAAAAACAGGA	7376
Qy	3470	YAGCTCTRGGAGTCTTTACACAGRTCCRAGGGYAGGCTTGCAACCTTGGCCTTGGCTTGA	3529
Db	7377	TAGCTCTAGAGTCTTTACACAGATCCGAGGGATGAGCTTGCAACCTTGGCCTTGGCTTGA	7436
Qy	3530	STAAAGGAAAYTGATGTAGTGCAAAAGGGTT	3559
Db	7437	CTAAGGAAATTGATGTAGTGCAAAAGGGTT	7466

RESULT	11
ADA02882	
ID	ADA02882 standard; DNA; 22436 BP.
XX	
XX	ADA02882;
XX	
DT	06-NOV-2003 (first entry)
XX	
DE	Human PAP carcinoma associated gene, SEQ ID NO:1400.
XX	
KW	Human; carcinoma associated; oncogene; carcinoma; cancer; breast;
KW	prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
KW	gene; ds.
XX	
OS	Homo sapiens.
XX	

PN	WO2003057146-A2.
XX	
PD	17-JUL-2003.
XX	
PF	26-DEC-2002; 2002WO-US041414.
XX	
PR	26-DEC-2001; 2001US-00035832.
XX	(SAGR-) SAGRES DISCOVERY.
PA	
XX	Morris DW;
PI	
XX	WPI; 2003-587068/55.
DR	
XX	
PT	New recombinant nucleic acid encoding carcinoma associated protein,
PT	useful for preparing compositions for treating carcinomas.
XX	
PS	Claim 1; SEQ ID NO 1400; 245pp; English.
XX	
CC	The invention relates to recombinant carcinoma associated (CA) nucleic
CC	acid sequences from mouse and human (ADA01482-ADA03094), and to
CC	recombinant carcinoma associated proteins (CAP) encoded by them. The
CC	invention also encompasses expression vectors and host cells comprising a
CC	CA nucleic acid, a polypeptide (especially an antibody) that specifically
CC	binds to the protein, and a biochip comprising CA nucleic acid or
CC	fragments thereof. The sequences of the invention were identified using
CC	oncogenic retroviruses, which insert into the genome of the host organism
CC	at random. Many of these do not carry transduced host oncogenes or
CC	pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC	direct consequence of the effects of proviral integration into host
CC	protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC	carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC	leukaemia) or a propensity to carcinoma by determination of the sequence
CC	of a CA gene, or by determination of CA gene expression in particular
CC	tissues. CA nucleic acids, proteins and antibodies are also useful as
CC	therapeutic agents and in screening and evaluating drug candidates. The
CC	present sequence represents a specifically claimed human CA nucleic acid
CC	sequence of the invention. Note: The complete sequence data for this
CC	patent did not form part of the printed specification, but was obtained
CC	in electronic format directly from WIPO at
CC	<a href="http://ftp.wipo.int/pub/published/pct/sequences">ftp.wipo.int/pub/published/pct/sequences</a> .

Seq	Sequence	22436 BP;	7692 A;	4872 C;	4296 G;	5576 T;	0 U;	0 Other;
Query Match	40.7%;	Score	3049.8;	DB	9;	Length	22436;	
Best Local Similarity	86.5%;	Pred. No.	0;					
Matches	3556;	Conservative	142;	Mismatches	320;	Indels	93;	Gaps 31;
QY	606	GATGGGAAACGTTCCCGCGNAGACAAAAACGCCCTTAAGACGTATTCTG	GAPAAATTGGGA	665				
DB	1581	GSTAGCAAAACATTCCCTCTCAAGCCAAAAAACACCCGAAGATGATTCTG	GAGAAATCGGGA	1640				
QY	666	MCAATTTTGACCCCTCAGACACTTAAGAAAGAAACACATTATTTCTTC	GCAGTGC	725				
DB	1641	CCAATTTTGACCCCTCAGACCGCTGAGAAAGAAATGACATTATTTCTT	CGCAGTATTCGCTG	1700				
QY	726	GC-----	ACTCTGAGGGAAGTATAAATTATAA	753				
DB	1701	GCCACAATATCTCTTCTGTGGGGGAGAAACGTGGCCTCTCTGAGGAAGTATAA	ATTATGA	1760				
QY	754	CACCATCTTACAGCTAGACVTCCTTTGTAGAA---	AAGGCAAAATGGAGTGAAGTGC	810				
DB	1761	CACCATCTTACAGCTAGACCTCTTTTGTAGAAAGAAAGGCAATGGAGTGA	AGCGCCATA	1820				
QY	811	AGTACAAACCTTCTTTTTCATTAAAGACAACTCACAAATTATGTAAAAAGTGT	GATTTATG	870				
DB	1821	TGTAACAATTTTCTTTTTCATTAAAGACAACTCGCTATTATGTAAAAAGTGT	GATTTATC	1880				
QY	871	CCCTACAGGAAGCCTTCAGAGTCTACCTCCCTATCCGAGCAT---	CCCCGACTCCTTCCCC	928				
DB	1881	CCCTACAAGAAGCCCTCAGAGTGTACCTTCCTACCCGAGCATCCCCCTACT	CTCTTCCCC	1940				
QY	929	AMYTATTAAGGACCCCCCTTTCACCCCAAAATGGTCCAAAGGAGATACACA	AAAGGGTAAA	988				



```
QY 3144 TTATCCAAAGGACCAARGGCCCTCAGTCAGGAAVRYATCCAGCCTATACCTGCTTATCCT 3203
Db 4130 TTATCCAAAGGTACCAGGGCCCTCAGTCGGGAACAATATGCGCCTATATCTGCTTGTCTT 4189
QY 3204 CATCYCAAAACCCCTAAAGCAACTAAGRRRTTCCCTTGGCRTAAYAGVTTCTGCCGAWA 3263
Db 4190 CATCCAAACCCCT--AAGCACTAAGAGGGTTCCTTGGCATACAGGTTTCTGCCAATA 4248
QY 3264 TGGATTCCCCAGGTGTGGCRAAATAGCCAGGYCATTAWATACASTAAATTAAGGAACTCA 3323
Db 4249 TGGATT--CCAGGTATGCAAAATAGCCAGATCATTTATATACATAAATTAAGGAACTCA 4307
QY 3324 GAAAGCCAAATCCCATTTARTAAATGAGYAMCTGAAGYMRAGTGGCTTCCAGGCCCC 3383
Db 4308 GAAAGCCAAATCCCATTTAGTAAGGTGGACACCTGAAGCACAAGCAGCTTCCAGGCTCT 4367
QY 3384 TAAAGAGGCCCTTAAACCCAGYCCAGTGTAAAGVTTGCCAACRGGGCAAGACTTTTST 3443
Db 4368 AAAGAA--GGCTCTAAACCCAGTTCAGTGTACGCTTGCCAAAGGGGCAAGATTTTCT 4425
QY 3444 TYATAYRTACAGAAAAAACAAGAAAYAGCTCTRGAGTCCCTTACACAGRTCCRAAGGAY 3503
Db 4426 TTAAATGTACAGAAAAAACAAGAAATAGCTCTAGAGTCCCTTACACAGGTCCGAGGAT 4485
QY 3504 GAGCTTGCAACCYRTGGCRYACCTGAATAAGGAAAYTGATGTAGTGGCAAAAGGTTGRCY 3563
Db 4486 GAACTTGCCAACTGTGGCATACCTCAGTAAGGAAATGTATGAGTGGCAAAAGGTTGGCC 4545
QY 3564 TCATTGTTTAYGGGTAGTGGCAGTAGCAGTYTAGTATCTGAAGCAGTTAAATAAT 3623
Db 4546 TCATTGTTTAYGGGTAGTGGCAGCGGTAGCAGTCTTAGTATCTGAAGCAGTTAAATAAT 4605
QY 3624 ACAGGAGAGATCTTACTGTGTGACATCTCATGAKGTGAAYRGCACTACTCAGTCTAA 3683
Db 4606 ACATGAGAGATCTTACTGTGTGACATCTCATGATGTGAACGGCACTACTCAGTCTAA 4665
QY 3684 AGGAGACTTGTGGCTGTGACAGCAACVGTTTACTTAAATRTCAGGCTCTATTACTTGAAG 3743
Db 4666 AGGAGACTTGTGGCTGTGACAGCATCGTTTACTTAAACATCATCAGGCTCTATTACTTGAAG 4725
QY 3744 GCAGTGTGCRACGTGCACTGTGCACTCTTAACCCAGYCNCACTTCTCCAGACAA 3803
Db 4726 GCCTGTGTGCAACTGTGCGGCTTTGCAACTCTTAAACCCAGCACTTCTTCCAGACAA 4785
QY 3804 TGAAGAAAAATARAAYATACTGTCAAACAAATTAATTTCTCAAACCTATGCCACCTCGAGG 3863
Db 4786 TGAAGAAAGATAGAACAATACTGTCAACAAGTAATTTGCTCAACCTACCCGCTCGAGG 4845
QY 3864 GGACCTTYPAGAGTTCCTVTTGACTGATCCYGACCTTCAACTTGTATATCTGATGGAAGTT 3923
Db 4846 GGACCTTTTAGAGTTCCCTGACTGATTCGACCTT--AACTTGTATATCTGATGGAAGTT 4904
QY 3924 CTTTGTAGAAAAAGCACTTCGAAAGYGGGTATGCACTGCTCAGTGTATGTAATGAATAY 3983
Db 4905 CTTTGTAGAAAAAGCACTTCGAA--AGTGTGTCAGGGATATATGGAATAC 4952
QY 3984 TTGAAAGTAATCCCTCACTCCAGGAACTAGTGTCTGCTGCAAGCACTAATAGCCYCA 4043
Db 4953 TTGAAATANTCCCTCACTCCAGGACTAGTGTCTGCTCAGCCGCAATACTAATAGCCCTCA 5012
QY 4044 YTKGGGCACTAGAAATTAGAGAGAAAGAAAGGYAAATATATATACAGACTCTRTATATG 4103
Db 5013 CTCGAGCACTAGAAATTAGAGAGAAAGAAAGGYAAATATATATATATAGACTCTAAGATATG 5072
QY 4104 CTYACTAGTONTCCATGCCATMRGCAATATGSAAGAAAGGGAATTCCTAACTTCVG 4163
Db 5073 CTTACTAGTCTCTCATGCCATGCAAGCAATATGAGAGAAAGGGAATTCCTAACTTCVG 5132
QY 4164 AGRGAACACTTATCAMACATCAGGAAGCCATTAGGARATTTATTATYTGCGWGTACAGAAC 4223
Db 5133 AGGGAACACTTATCAACCATCAGGAAGCCATTAGGAGATTTATTATYTGCGTGTACAGAAC 5192
```

```
QY 4224 CTABAGAGTGGMAGTCTTACACATGCYGGGCTCATANAAGGAAGAAAGGAAGGAATAS 4283
Db 5193 CCAAGAGGTGGCAGTCTTACACTGCTAGGTCTATCAGAGAAGAAAGGAAGGAATAG 5252
QY 4284 AAGGAAYTGCAAGCAKATATTGAAGCMAAAAGAGCTGCAAGGCAGGACCCCTCCATTAG 4343
Db 5253 AAGGGAACCAAGCAGATATTGAGCCAAAAGAGCTGCAAGACAGGACCCCTCCATCAG 5312
QY 4344 AAATGCTTATTAACCTTCCCTTAGTATAGGTAAATCCCTTCCGGGAAACCAAGCCCACT 4403
Db 5313 AAATGCTTATAGAGGACCCCTAGTGTGGGTAACCCCTCCAGGAAACCAAGCCCACT 5372
QY 4404 ACTCAGCAGGAGAAACAGAAATGGGGAACCTCACAGG--CAGTTTTCTCCCTCGGAC 4460
Db 5373 ACTCAGCAGGATGAATAGAATGGGGAACCTCATGAGGACATACCTTCTCCCTCAGGAT 5432
QY 4461 GGTTAGCCACTGAAGAAGGGAATACTTTTGCCTGCAACTATCCAAATGGAATTTACTTTA 4520
Db 5433 GGCTAGCCACCAAGAGAGAAATACTTTTGCCTGCACTAACCAATGGAATTTACTTTA 5492
QY 4521 AAACCTTTCATCAAACTTTTCACTTAGGCATCGATAGCACCCATCARATGGCCAAATCAT 4580
Db 5493 AAACCTTTCACAAAACCTTTCACTTAGACATTTGATAGCACCCATCAGATGGCCAAATTTAT 5552
QY 4581 TATTTACTGGACAGGCTTTTCAAACTATCAAGCABATATCAAGGCGCTGTGAAGTGT 4640
Db 5553 TATTTACTGGACAGGCTTTTCAAACTATCAAGCAGATAGTCAGGCGCTGTGAAGTGT 5612
QY 4641 GCCARAAATAATCCCTGCTTATCGCCA 4671
Db 5613 GCCAAGAAATAACCCCTGCACTCGAGCCA 5643

RESULT 12
ADB72620
ID ADB72620 standard; DNA; 22436 BP.
XX
AC ADB72620;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human PAP gene.
XX
KW human; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX
OS Homo sapiens.
XX
PN WO2003008583-A2.
XX
PD 30-JAN-2003.
XX
PF 26-DEC-2001; 2001WO-US051291.
XX
PR 02-MAR-2001; 2001US-00798586.
PR 23-OCT-2001; 2001US-00004113.
PR 08-NOV-2001; 2001US-00052482.
PR 30-NOV-2001; 2001US-00997722.
PR 20-DEC-2001; 2001US-00034650.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW, Engelhard EK;
XX
WPI; 2003-239337/23.
XX
PT New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
PT cancers, neoplasm, adenocarcinoma, or sarcomas.
XX
PS Claim 1; SEQ ID NO 448; 2304pp; English.
XX
CC The invention relates to a novel recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the 660 sequences fully defined
```

in the specification. A polynucleotide of the invention has cytostatic activity, and may have a use in gene therapy, or in a vaccine. The recombinant nucleic acids and polypeptides are useful for treating carcinomas, e.g. lymphomas, cancers, neoplasms, adenocarcinoma, and sarcomas. The present sequence represents a human gene of the invention.

Query Match	40.7%;	Score 3049.8;	DB 10;	Length 22436;
Best Local Similarity	86.5%;	Pred. No. 0;		
Matches 3556;	Conservative 142;	Mismatches 320;	Indels 93;	Gaps 31;
Qy	606	GATGGGAAACGTTCCCGCAGACAAAAAAGCCCTAAGAGTATTTCTGGARAATTCGGAA	665	
Db	1581	GGTAGCAACATTCCTCTCAGGCCAAAAAACCCCGAAGATGTATTTCTGGAGAATTCGGGA	1640	
Qy	666	MCAATTTGACCTTCAGACACTAAGAAAGAAACGACTTATATTTCTTCGCAGTGCGCCCTG	725	
Db	1641	CCAATTTGACCTTCAGACGCTGAGAAAGAAATGACTTATATTTCTTCGCAGTATTGCCCTG	1700	
Qy	726	GC-----	753	
Db	1701	GCACAATATCTCTTCATGCGGGGAGAAACGTGGCCCTCTTGAGGGAAGTATAAATATTAGA	1760	
Qy	754	CACCATCTTACAGCTAGACACTTCTTTCTGTAGAAA---AAGGCAAAATGGAGTGAAGTGCCATA	810	
Db	1761	CACCATCTTACAGCTAGACCTCTTTTCTGTAGAAAGAGGCNAATGGAGTGNAGGCCCATTA	1820	
Qy	811	AGTACAAACTTTCTTTTTCATTTAAGAGACAACACTCAAAATATGTATAAAAAGTGTGATTTATG	870	
Db	1821	TGTACAAATTTTCTTTTCATTTAAGAGACAACCTGCTATTATGTATAAAAAGTGTGATTTATC	1880	
Qy	871	CCCTACAGGAAGCCTTCAGAGTGTACCTCCCTATCCAGCAT--CCCGACTCTCTTCCCC	928	
Db	1881	CCCTACAAGAAGCCCTCAGAGTGTACCTTCTTACCCAGCATCCCCCTACTCTTCTCCCC	1940	
Qy	929	AMTATAAAGGACCCCTTCAACCCAAATGGTCCAAAGGAGATACACAAAAGGGTAAA	988	
Db	1941	AACCTATAAGGACCCCTTCAACCCAAATAGTCCNAAGAGAGATAGACAAAAGGGTAAA	2000	
Qy	989	CAGTGAACCAAAGAGTGCCTAATTTCCCAATATATGACCCCTCCCAAGCAGTGGGAGAA	1048	
Db	2001	CAATGAATCAAAGAGTGCCTAATTTCTCCGATATATGCCACT--CCAAACAGTGGGAGAG	2059	
Qy	1049	GAGAAATTCGGCCAGCCAGAGTGCATGTCGCTTTTCTCCAGACTTAAAGCAATATA	1108	
Db	2060	GAGAAATTCGGCCAGCCAGAGTGCATGATGATCAATTTCTCTCTCAGTCTTAAAGCAATATG	2119	
Qy	1109	AAACAGACTTTAGGTAAATTTCTCAGATAAYCCTGATGGCTATATTTGRTGTTTTTCAAGGGT	1168	
Db	2120	AAATAGACCTTAGTAAATTTCTCAGCTAACCTGATGGCTATATTTGATTTTTTACAGGGT	2179	
Qy	1169	TAGACAATTTCTTTGATCTGACATGGAGAGATATATATGTCACTGCTPAAATCAGACACTA	1228	
Db	2180	TAGACAATTCCTTTGATCTGACATGGAGAGATATA-ATGTTACTGTTTAAATCAGACACTA	2238	
Qy	1229	ACCCAAATCAGAGAGTGCACCATAACTGCAGCCCTGAGRGTTTGGCGATCTCTGGTAT	1288	
Db	2239	ACCCAAATCAGAGAGTGCAGCCCATAACTGCAGCCCGAGAGTTTGGCGATCTCTGGTAT	2298	
Qy	1289	CTCAGTCAGGTCAATGGATANGATGACACAGAAAGGAANAATGATTTCCCCACAGGC	1348	
Db	2299	CTCAGTCAGGTCCAT--CATAGGATGAAAACAG-AGGAAGAGAGACGATTTCCCCACAGGC	2355	
Qy	1349	CAGCARGCAGTTCGCAGTCTTASACCTTCATTTGGGGACAGAGAAATCATGTAAATGGGAGA	1408	
Db	2356	CAGCARGCAGTTCGCAGTGTAGACCTTCATAGGG--ACACAGAAATCAGAAACAT-GGAGA	2411	
Qy	1409	TTGGTGTGCAGACATTTTGTCTTAACTTGTGTCTASAGGACTAAGGAAACTAGGAAGA	1468	
Db	2412	ATGGTGACGACACATTTTGTCTTAACTTGTGTCTAGGAAGGACTAAGGAAACTAGGAAG-A	2470	
Qy	1469	ARTCTAYGAATTACTCAATGATGTCCACCATAACACAGGGGGAAGGAAAGAAATCTCTACT	1528	

[illegible]



Db 3540 GCCCAACAGGACTCCAAAGATTTGTTAAGGACCGAAAAGCCCAAGGCTCTAGTAAACCA 3599  
Qy 2604 TGCATAAATCCCTGCGAGTAATTCCTGAGTGGATTGAGGAGGCACAGAAACCCAGTGGACA 2663  
Db 3600 TGCATAAGCCCTTGCAATATCC-----AATTTTAGGAGTACAGAAACCCAGTGGACA 3652  
Qy 2664 GTGGAGGGTTAGTGCAGAGATCTCAGAGATTATCAATGGAGGCGTTGTCTCTTTATACCCA 2723  
Db 3653 GTGGA-GGTTAGTGCAGAGATCTCAGAGATTATCGAT-TAGGCTGTGTGTTCTCTATATCA 3710  
Qy 2724 GCTGACCTAGCCCTTATATCTGTGMYTTCCCAAATACCAGAGAAAGCAGAGTGGTTTACA 2783  
Db 3711 GCTGTACCTAACCCCTTATACCTGCTTTCCCAAATACCAGAGAAAGCAGAGTGGTTTACA 3770  
Qy 2784 STCCGTGACCTTMAGAGTCCCTTCTCTGATCCCTGTACATCCCTGTACATCTCAATCTTGTG 2843  
Db 3771 GTCTCTGAGCTTTAAGGATGTCTTTTCTGTCATCCCTGTACATCTCTCAATCTTTTA 3830  
Qy 2844 TTTGCTTTGAGATACTTCAAAACCCARCATCTCAACTCACCTGACCTGTTTACCCCA 2903  
Db 3831 TTTGCTTTGAGATCTTTCGAACCCATGTCTCACTCACCTGGACTGTTTACCCAG 3890  
Qy 2904 GGGTTCCAGGATAGYCCCCATCTATTTGGCCAGGCAATTAGCCCAAGACTTGAGYCAATYM 2963  
Db 3891 GGGTTCCAGATAGCCCCCATCTATTTGGCCAGGCAATTAGCCCAAGACTTGAGCCAAATTC 3950  
Qy 2964 TCATACCTGGACACTTGTCTCTTCRGTAKGTGGATGATTTACTTTTTRGCGCCVRRITCA 3023  
Db 3951 TCATACCTGGACACTTGTCTCTTGGTACATGATGATTTACTTTTAGCCACCCATTTCA 4010  
Qy 3024 GAAACCTTGTGCATCAAGCCACCAAGCRCTCTTMAATTTCTCGVACCTGTGGCTAC 3083  
Db 4011 GAAGCTTTGTGCATCAAGCCACCAAGCACTCTTAAATTTCTCGTACCTGTGGCTAC 4070  
Qy 3084 AMWGTTTTCAAAASARAGCTCARCTCTGCTCACAGAGGTTTAAATACTTAGGCTAARA 3143  
Db 4071 AA-GATTCANACCAAGGCTCAGCTCTGCTCACAGAGTTTAAATACTTAGGCTAANA 4129  
Qy 3144 TTATCAAAGGCACCAAGGCCCTCAGTGGAGAAATATCCAGCTTATCTGCTTATCTCT 3203  
Db 4130 TTATCAAAGGTACCAAGGCCCTCAGTGGGGAACATATCGACCTTATCTGCTTGTCT 4189  
Qy 3204 CATYCAAACCTTAAGCAACTAAGRPRTCTTGGORTAAYAGGYTTCTGCCGAWA 3263  
Db 4190 CATCCAAAACCTT-AAGCAACTAAGAGGGTTCTTGGCATAAACAGGTTTCTGCCAATA 4248  
Qy 3264 TGGATTCCCGAGTWTGGRAAATAGCCAGGYCATTAATATACATACTAATTAAGGAACTCA 3323  
Db 4249 TGGATT-CCAGGTATGGCAAAATAGCCAGATCATATATACATTAATTAAGGAACTCA 4307  
Qy 3324 GAAAGCCAAATACCAATTTARTAAGATGGAYAMCTGAAGYMAAGTGGCTTTCCAGGCCCC 3383  
Db 4308 GAAAGCCAAATACCAATTTAGTAAAGTGGACACCTGAAGCACAAGCAGCTTTCCAGGCTCT 4367  
Qy 3384 TAAAGAGCCCTTAACCCCAAGYCCAGTGTAAAGYTTGCCAACCGGCGAGACTTTTST 3443  
Db 4368 AAAGAA--GGCTCTAACCCCAAGTCCAGTGTACGCTTGCCAAAGGGGCAAGATTTTCT 4425  
Qy 3444 TYATAYRTACAGAAAAAACAAGAAAYAGCTCTCGAGTTCCTTACACAGRTCCRAGGGAY 3503  
Db 4426 TTAATGTTCACAGAAAAAACAAGAAATAGCTCTAGAGTTCCTTACACAGGTCGAGGGAT 4485  
Qy 3504 GAGCTTGCAACYRTGGCRYACTGASTAAGAAAYTGATGTAGTGGCAAGGGTTGRCY 3563  
Db 4486 GAACTTGCAACTGTGGCATACCTGATTAAGGAAATTTGATGAGTGGCAAGGGTTGGCC 4545  
Qy 3564 TCATTGTTTAYGGTAGTGGTCAGTACGATGYKTATGATCTGAGCAGTAAATATAT 3623  
Db 4546 TCATTGTTTATGGGTAGTGGCAGCGGTAGCAGTCTTAGTATCTGAAGCAGTTAAATATAT 4605  
Qy 3624 ACAGGGRAGAGATCTTACTGTGTGGACATCTCATGAKGTGAAYRGCATCTCACTGCTAA 3683  
Db 4606 ACATGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAACGGCAGTCACTCACTGCTAA 4665

Qy 3684 AGGAGACTTGTGGCTGTCCAGACAAACYGTTTACTTTAAATRTCCAGGCTCTATTACTTTGAARG 3743  
Db 4666 AGGAGACTTGTGGCTGTCCAGACGATGTTTACTTTAAACATCAGGCTCTATTACTTTGAARG 4725  
Qy 3744 GCCAGTGTGCRATCTGTGCAACTTGTAAACCAGYCNCAATTTCTTCCAGACAA 3803  
Db 4726 GCCTGTGTGCAACTGTGGCGCTTTGCAACTCTTAAACCAGCCACATTTCTTCCAGACAA 4785  
Qy 3804 TGAAGAAAGATARAAYATAAATCTGTCAACAARTAAATTTCTCAAACTATGCCACTCGAGG 3863  
Db 4786 TGAAGAAAGATAGAACAATACTGTCAACAAGTAATGCTCAAACTACCCGCTCGAGG 4845  
Qy 3864 GGACCTTGTAGAGTTCCTTGTGATCCYGCCTTCAACTTGTGTACTGTAGTGAAGTT 3923  
Db 4846 GGACCTTGTAGAGTTCCTTGTGATCTTCCGACTTCCGACTT-AACTGTGTATCTGATGAAGTT 4904  
Qy 3924 CTTTGTAGAAAAAGGACTTCGAAAAAGYGGGTATGCAGTGCTAGTGATATGAATATG 3983  
Db 4905 CTTTGTAGAAAAAGGACTTGA--AGTGTGAGGATTAATGGAATAC 4952  
Qy 3984 TTGAAGTAATCCCTCTCACTCCAGAACTAGTGTCTAGTGTGCAAGAACTAATAGCTTCA 4043  
Db 4953 TTGAATAATAATCCCTCTCACTCCAGAACTAGTGTCTAGCCGCAATACTAATAGCCCTCA 5012  
Qy 4044 YTKGGCACTAGATATAGGAGAAAGAAAGGAAATATATATACAGACTCTTARTATG 4103  
Db 5013 CTCGAGCACTAGATATAGGAGAAAGAAAGGTAATATATACATAGACTCTAAGTATG 5072  
Qy 4104 CTYACCTAGTCNTCCATGCCCATMRGCAATATGSARAGAAAGGAAATCTTAACTTCVG 4163  
Db 5073 CTTACCTAGTCTCCATGCCCATGCAGCAATATGGAGAAAGGAAATCTTAACTTCTG 5132  
Qy 4164 AGRGAACACTTATCAMACATCAGGAAGCCATTAGGARATTTATYTGCGWGTACAGAAAC 4223  
Db 5133 AGGGAACACTTATCAACCATCAGGAAGCCATTAGGAGATTTATTTGGCTGTACAGAAAC 5192  
Qy 4224 CTARAGAGTGTGMACTTCTTACCTGCTGGGTCTCATANAAGAAAGAAAGGAAAGTAS 4283  
Db 5193 CCNAAGAGTGTGGAGTCTTACCTGCTAGGCTCATAGAGAAAGAAAGGAAAGTATG 5252  
Qy 4284 AAGRNAATGCAAGCAKATATGAAGCMAAAGAGCTGCAAGGAGAGACCTCCATTAG 4343  
Db 5253 AAGGAACCAACAGCAGATATTTAGCCAAAGAGCTGCAAGACAGGACCTCCATTAG 5312  
Qy 4344 AAATGCTTTATTAACCTTCCCTTAGTATAGGTAATCCCTTCCGGGAAACCAAGCCCCAGT 4403  
Db 5313 AAATGCTTTATAGAGGACCCCTAGTGTGGGTAAACCCCTCCAGGAACCAAGCCCCAGT 5372  
Qy 4404 ACTCAGCAGGAGAAACAGANTGGGAACTCACAGG--CAGTTTTCTCCCTCGGGAC 4460  
Db 5373 ACTCAGCAGGATGATAGATGGGAACTCATAGGAGACATCTTTCTCCCTCGAGGAT 5432  
Qy 4461 GGTAGCCACTGAAGAGGAAATATCTTTGCTGCAACTATCCAATGGAAATTTACTTTA 4520  
Db 5433 GGTAGCCACCAAGAGGAAATATCTTTGCTGCACTAACCAATGGAAATTTACTTTA 5492  
Qy 4521 AAAACCTTTCATCAACCTTTCACTTAGGCATCGATAGCACCATCARATGGCCAAATCAT 4580  
Db 5493 AAAACCTTTCACCAAACTTTCACTTAGACATTTGATAGCACCATCAGATGGCCAAATCAT 5552  
Qy 4581 TATTACTGGACAGGCTTTTCAAACTATCAAGCABATATCTAGGCGCTGTGAATGT 4640  
Db 5553 TATTACTGGACAGGCTTTTCAAACTATCAAGCABATATCTAGGCGCTGTGAATGT 5612  
Qy 4641 GCCARAAAAATATCCCTGCTCTATCGCCA 4671  
Db 5613 GCCAAGAAATAACCCCTGCACTGCAGGCCA 5643

RESULT 13  
ADC85361  
ID ADC85361 standard; DNA; 22436 BP.







Db 5193 CCAAGAGGTGGCAGTCTTACACTGTAGGTCATCAGAGAAAGGAAGGAATAG 5252  
 QY 4284 AAGGAAATGCCAAGCAKATATTGAACMAAAGAGCTGCAAGCGAGGACCTCCATTAG 4343  
 Db 5253 AAGGAAATGCCAAGCAKATATTGAAGCCAAAGAGCTGCAAGCGAGGACCTCCATTAG 5312  
 QY 4344 AAATGCTTATTAACCTTCCCTTGTATAGGTATCCCTTCCGGAACCAAGCCCGAGT 4403  
 Db 5313 AAATGCTTATAGAGGACCTTGTGTGGGTAACTCCCTTCCGGAACCAAGCCCGAGT 5372  
 QY 4404 ACTCAGCAGGAGAAACAGAAATGGGAACTTCAAGAGG--CAGTTTTCTCCCTCGGGAC 4460  
 Db 5373 ACTCAGCAGGATGAATAGATGGGAACTTATGAGGACATATCTTCTCCCTGAGGAT 5432  
 QY 4461 GGTAGGCACTGAAGAGGAAATATCTTTGCTGCAATCTCAATGGAATTA 4520  
 Db 5433 GGCTAGCACCACCAAGAGGAAATATCTTTGCTGCAATCTCAATGGAATTA 5492  
 QY 4521 AAACCTTTCATCAACCTTTTCACTAGGATGATAGGACCATCAATGGAATTA 4580  
 Db 5493 AAACCTTTCATCAACCTTTTCACTAGGATGATAGGACCATCAATGGAATTA 5552  
 QY 4581 TATTACTGACGAGGCTTTTCAAACTATCAAGCAATATCAAGGCTGTGAATGT 4640  
 Db 5553 TATTACTGACGAGGCTTTTCAAACTATCAAGCAATATCAAGGCTGTGAATGT 5612  
 QY 4641 GCCAARAATAATCCCTGCTTATGCGCA 4671  
 Db 5613 GCCAARAATAATCCCTGCTTATGCGCA 5643

RESULT 14

ADM74477  
 ID ADM74477 standard; DNA; 22436 BP.  
 AC ADM74477;  
 XX  
 DT 01-JUL-2004 (first entry)  
 XX  
 DE Human carcinoma associated (CA) nucleic acid #73.  
 XX  
 KW Human: carcinoma associated nucleic acid; CA nucleic acid; gene; ds;  
 KW carcinoma associated protein; CAP; carcinoma; leukaemia; lymphoma;  
 KW cytostatic.  
 XX  
 OS Homo sapiens.  
 XX  
 XX US2004072154-A1.  
 XX  
 PD 15-APR-2004.  
 XX  
 PF 30-NOV-2001; 2001US-00997722.  
 XX  
 PR 22-DEC-2000; 2000US-00747377.  
 PR 02-MAR-2001; 2001US-00798586.  
 XX  
 PA (MORRIS D W.  
 PA (ENGELHARD E K.  
 XX  
 PI Morris DW, Engelhard EK;  
 XX  
 DR WPI; 2004-328562/30.  
 XX  
 PT New carcinoma associated gene or protein, useful for preparing a  
 PT composition for diagnosing or treating carcinoma e.g., leukemia or  
 PT lymphoma.  
 XX  
 PS Claim 1; SEQ ID NO 148; 29pp; English.  
 XX  
 CC The invention relates to new recombinant nucleic acids. The invention  
 CC also relates to a host cell comprising a recombinant nucleic acid or  
 CC expression vector, an expression vector comprising a recombinant nucleic  
 CC acid, a recombinant protein, a method of screening for drug candidates, a

CC method of screening for a bioactive agent capable of binding to a  
 CC carcinoma associated protein (CAP) encoded by a nucleotide sequence, a  
 CC method of screening for a bioactive agent capable of modulating the  
 CC activity of a CAP, a method of evaluating the effect of a candidate  
 CC carcinoma drug, a method of diagnosing carcinomas, a method for inhibiting  
 CC the activity of a CAP, a method of treating carcinomas, a method of  
 CC neutralising the effect of a CAP and a method of diagnosing carcinoma or  
 CC propensity to carcinoma. A method of evaluating the effect of a candidate  
 CC carcinoma drug comprises administering the drug to a patient, removing a  
 CC cell sample from the patient and determining alterations in the  
 CC expression or activation of a gene comprising the nucleotide sequence. A  
 CC method of diagnosing carcinoma comprises determining the expression of  
 CC one or more genes comprising the nucleic acid sequence in a first tissue  
 CC type of a first individual and comparing the expression of the gene from  
 CC a second normal tissue type from the first individual or a second  
 CC unaffected individual, where a difference in the expression indicates  
 CC that the first individual has carcinoma. A method of inhibiting the  
 CC activity of a CAP comprises binding an inhibitor to the CAP. Treating  
 CC carcinomas comprises administering to a patient an inhibitor of CAP.  
 CC Neutralising the effect of a CAP comprises contacting an agent specific  
 CC for the CAP. The polypeptide specifically binds to the protein encoded by  
 CC the nucleic acid. It comprises an antibody that specifically binds to the  
 CC protein encoded by the nucleic acid. The nucleic acids are useful for  
 CC preparing a composition for diagnosing or treating carcinoma e.g.,  
 CC leukaemia or lymphoma. This sequence represents a human carcinoma  
 CC associated (CA) nucleic acid of the invention. Note: The sequence data  
 CC for this patent did not form part of the printed specification but was  
 CC obtained in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html.

XX  
 SQ Sequence 22436 BP; 7692 A; 4872 C; 4296 G; 5576 T; 0 U; 0 Other;  
 Query Match 40.7%; Score 3049.8; DB 12; Length 22436;  
 Best Local Similarity 86.5%; Pred. No. 0;  
 Matches 3556; Conservative 142; Mismatch 320; Indels 93; Gaps 31;  
 QY 606 GATGGAAACGTTCCCGCAAGCAAAACGCGCTTATCTGAGACGTTATCTGGAATTTGGGA 665  
 Db 1581 GGTAGCAAAACATTCCTCTCAAGGCAAAACACCCGAGATGTTATCTGAGATCGGGA 1640  
 QY 666 MCAATTTGACCTCAGACACTAAGAAAGAAAGACATTTATTTCTTCTGAGTCCGCTG 725  
 Db 1641 CCAATTTGACCTCAGACACTAAGAAAGAAAGACATTTATTTCTTCTGAGTATTCCTG 1700  
 QY 726 GC-----ACTCCTGAGGAGATATAAATATAA 753  
 Db 1701 GCACATATATCTCTTATGAGGAGAAACGTTGAGGCTCTGAGGAGATATAAATATAA 1760  
 QY 754 CACCATCTTACAGTAGACTCTTTTGTAGAA---AAGGCAATGAGTGAAGTGCATA 810  
 Db 1761 CACCATCTTACAGTAGACTCTTTTGTAGAAAGAGCAATGAGTGAAGTGCATA 1820  
 QY 811 AGTACAACTTTCTTTTCAATTAAGAGACAACTCAAAATATGTAAAAAGTGTATTATG 870  
 Db 1821 TGTACAAATTTCTTTTCAATTAAGAGACAACTCGCTATATGTAAAAAGTGTATTATC 1880  
 QY 871 CCTACAGAGAGCTTCAAGTCTACCTCCCTATCCAGCAT--CCCCAGCTCTTCCCC 928  
 Db 1881 CCTACAGAGAGCTTCAAGTCTACCTCCCTATCCAGCAT--CCCCAGCTCTTCCCC 1940  
 QY 929 AMYTAATAAGGACCCCTTCAACCAATGTGCAAAAGGAGATAGACAAAAGGTAAA 988  
 Db 1941 AACTTAATAGGACCCCTTCAACCAATGTGCAAAAGGAGATAGACAAAAGGTAAA 2000  
 QY 989 CAGTGAACCAAGAGTGCATATTTCCCAATTTATGACCCCTCCCAAGCAGTGGAGGAA 1048  
 Db 2001 CAATGAATCAAGAGTGCATATTTCTCGATTATGCCCCACT--CCAAAAGTGGAGGAG 2059  
 QY 1049 GAGATTTGGCCCGACCGAGTGCATGCTGCTTTTCTCCAGCTTAAGACAAATAA 1108  
 Db 2060 GAGATTTGGCCCGACCGAGTGCATGCTGCTTTTCTCTCAGTCTTAAAGCAAAATG 2119  
 QY 1109 AAACAGACTTAGGTAAATTTCTCAGATAAATCTGATGGCTATATTGTTGTTTTACAAGGT 1168











[illegible]

Search completed: August 1, 2006, 19:13:51  
Job time : 6567 secs

---